

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 9, 2002, 18:36:04 ; Search time 21.8726 Seconds
(without alignments)
955.720 Million cell updates/sec

Title: US-09-830-144-4

Perfect score: 2580

Sequence: 1 MAAQRRLSQSQPSWTDD.....AEFYRLMSVDHGQSVWTAP 504

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2580	100.0	504	1 TAB1_HUMAN	Q15750 homo sapien
2	190	7.4	300	1 P2C_PARTE	P49444 paramescium
3	175	6.8	399	1 P2C4_ARATH	P49598 arabidopsis
4	168	6.5	423	1 P2C2_ARATH	O04719 arabidopsis
5	167.5	6.5	434	1 P2C1_ARATH	P49597 arabidopsis
6	160.5	6.2	370	1 P2C3_SCHPO	Q09172 schizosacch
7	154	6.0	538	1 PDPI_RAT	O84483 rattus norv
8	151.5	5.9	491	1 P2C1_CABEL	P49595 caenorhabdi
9	148.5	5.8	356	1 P2C2_CABEL	P49596 caenorhabdi
10	147	5.7	529	1 PDPI_HUMAN	Q9p219 homo sapien
11	145	5.6	538	1 PDPI_BOVIN	P35816 bos taurus
12	143.5	5.6	414	1 P2C3_SCHPO	Q09173 schizosacch
13	143	5.5	538	1 PDPI_HUMAN	Q9p0j1 homo sapien
14	139.5	5.4	388	1 P2C3_ARATH	P49599 arabidopsis
15	134.5	5.2	383	1 P2C4_SCHPO	O14156 schizosacch
16	131.5	5.1	543	1 P2CG_BOVIN	P79126 bos taurus
17	128	5.0	382	1 P2CA_HUMAN	P35813 homo sapien
18	128	5.0	530	1 PDPI_RAT	O84484 rattus norv
19	127	4.9	382	1 P2CA_RABIT	P35814 oryctolagus
20	127	4.9	454	1 FEM2_HUMAN	P49593 homo sapien
21	126	4.9	382	1 P2CA_RAT	P20650 rattus norv
22	125.5	4.9	546	1 P2CG_HUMAN	O15355 homo sapien
23	125	4.8	382	1 P2CA_MOUSE	P49443 mus musculu
24	124	4.8	382	1 P2CA_BOVIN	O62829 bos taurus
25	122	4.7	347	1 P2C1_SCHPO	Q01571 schizosacch
26	121	4.7	2145	1 CYAA_PODAN	Q01513 podospora a
27	120	4.7	1692	1 CYAA_SCHPO	P14605 schizosacch
28	118	4.6	390	1 P2CB_MOUSE	P36993 mus musculu
29	118	4.6	542	1 P2CG_MOUSE	Q61074 mus musculu
30	118	4.6	2493	1 CYAA_USTMA	P49606 ustilago ma
31	117.5	4.6	393	1 P2CA_YEAST	P38089 saccharomyc
32	117	4.5	390	1 P2CB_RAT	P35815 rattus norv
33	117	4.5	406	1 P2C_LEICH	P36982 leishmania

RESULT 1

ID	TAB1_HUMAN	STANDARD;	PRT;	504 AA.
AC	Q15750;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	TAK1-binding protein 1.			
GN	TAB1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=96216294; PubMed=8638164;			
RA	Shibuya H., Yamaguchi K., Shirakabe K., Tonegawa A., Gotoh Y.,			
RA	Ueno N., Irie K., Nishida E., Matsumoto K.;			
RT	"TAB1: an activator of the TAK1 MAPKKK in TGF-beta signal			
RT	transduction.";			
RL	Science 272:1179-1182(1996).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20057165; PubMed=10591208;			
RA	Dunham I., Hunt A.R., Collins J.E., Bruskiewicz R., Beare D.M.,			
RA	Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,			
RA	Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,			
RA	Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,			
RA	Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,			
RA	Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,			
RA	Dhani P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,			
RA	Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,			
RA	Gilbert J.G.R., Goward M.E., Gratham D.V., Griffiths M.N.D., Hall C.,			
RA	Hall R.E., Hall-Ramlyn G., Heathcote R.W., Ho S., Holmes S.,			
RA	Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,			
RA	Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,			
RA	Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann O.T.,			
RA	Mcclay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,			
RA	Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,			
RA	Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,			
RA	Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,			
RA	Soderlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M.,			
RA	Vaughn M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,			
RA	Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,			
RA	Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,			
RA	Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,			
RA	Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuana S.,			
RA	Roe B.A., Chen F., Chu Y., Hu P., Hua A., Kenton S., Do A., Do T.,			
RA	Dorman A., Fang F., Fu Y., Hu P., Loh P., Malaj E., Nguyen T., Pan H.,			
RA	Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,			
RA	Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,			
RA	Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,			
RA	Zhan M., Zhang G., Chissoe S., Murray J., Miller N., Minx P.,			
RA	Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,			

ALIGNMENTS

RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
 RA Hinds K., Kemp K., Latreille P., Layman D., Ozerky P., Rohlfing T.,
 RA Scheet P., Walker C., Wamsley A., Wohldmann P., Peglin K., Nelson J.,
 RA Korf I., Bedell J.A., Hillier L., Mardis E., Waterston R., Wilson R.,
 RA Emanuel B.S., Shaikh T., Kurahashi H., Saita S., Budarf M.L.,
 RA McDermid H.E., Johnson A., Wong A.C., Morrow B.E., Edelmann L.,
 RA Kim U.J., Shizuya H., Simon M.I., Dunanski J.P., Peyrard K.P., Kedra D.,
 RA Serousi E., Fransson I., Tapia I., Bruder C.E., O'Brien K.P., Lane L.,
 RA Wilkinson P., Bodenreich A., Hartman K., Hu K., Khan A.S., Lane L.,
 RA Tiliun Y., Wright H.,
 RT "The DNA sequence of human chromosome 22."
 RL Nature 402:489-495(1999).
 CC -!- FUNCTION: MAY BE AN IMPORTANT SIGNALING INTERMEDIATE BETWEEN TGFB
 CC RECEPTORS AND MAP3K/TAK1.
 CC -!- SUBUNIT: Interacts with MAP3K7 and with BIRC7.
 CC -!- TISSUE SPECIFICITY: UBIQUITOUS.
 CC -!- SIMILARITY: CONTAINS 1 PP2C-LIKE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

CC EMBL: U49928; AAC12660.1; -
 DR EMBL: Z83845; CAB55304.1; -
 DR GeneW: HGNC:18157; MAP3K7IP1.
 DR MIM: 602615; -
 DR InterPro: IPR001932; PP2C-like.
 DR Pfam: PF00481; PP2C; 1.
 DR SMART: SM00332; PP2C; 1.
 DR DOMAIN 64 368 PP2C-LIKE.
 FT 452 457 POLY-SER.
 SQ SEQUENCE 504 AA; 54644 MW; A45743288718983A CRC64;

Query Match 100.0%; Score 2580; DB 1; Length 504;
 Best Local Similarity 100.0%; Pred. No. 3,1e-164;
 Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAQRSLQSQQSWTDLPLCHLSGSGASNRYSADGKGTSPEDSWLKFRSEN 60
 DB 1 MAQRSLQSQQSWTDLPLCHLSGSGASNRYSADGKGTSPEDSWLKFRSEN 60
 QY 61 NCFYGVNGYDGNRVTFVFAQRLSABELLGQLNAEHADVRVLLQAFDVERSFLES 120
 DB 61 NCFYGVNGYDGNRVTFVFAQRLSABELLGQLNAEHADVRVLLQAFDVERSFLES 120
 QY 121 IDDALEAKXSLSQSLPEGVPHQQLPROYOKILERLKTLEREISGAMVAVLNNKLYV 180
 DB 121 IDDALEAKXSLSQSLPEGVPHQQLPROYOKILERLKTLEREISGAMVAVLNNKLYV 180
 QY 181 ANVGNTRALLCKSTVDGLQVTLNDHTTENDELFRLSQGLDAGKIKOYGII CGQEST 240
 DB 181 ANVGNTRALLCKSTVDGLQVTLNDHTTENDELFRLSQGLDAGKIKOYGII CGQEST 240
 QY 241 RRIQYKVKYGYTDIDLISAASKPIIAPEIHGAQPLDGYTGLVLMSEGLYKALEAAH 300
 DB 241 RRIQYKVKYGYTDIDLISAASKPIIAPEIHGAQPLDGYTGLVLMSEGLYKALEAAH 300
 QY 301 GGGQANQELIAMIITDEFKQKTSLDVAVAQVDRVRIHSDTFPAGSGEGARCPRHEDMTL 360
 DB 301 GGGQANQELIAMIITDEFKQKTSLDVAVAQVDRVRIHSDTFPAGSGEGARCPRHEDMTL 360
 QY 361 LVRRNGYPLGEMSOPTPSPAPAGRVYVPSVPSAOSTSKTSYTLISLWPSQGMWNG 420
 DB 361 LVRRNGYPLGEMSOPTPSPAPAGRVYVPSVPSAOSTSKTSYTLISLWPSQGMWNG 420
 QY 421 AHSASTLDEARPTLTNOSPPTLTLOSTNTHTOSSSSSDGGLFRSRPAHSLPGEEDGRVPE 480
 DB 421 AHSASTLDEARPTLTNOSPPTLTLOSTNTHTOSSSSSDGGLFRSRPAHSLPGEEDGRVPE 480

QY 481 YVDEAFERYRLMSVDHGQSVVTAP 504
 DB 481 YVDEAFERYRLMSVDHGQSVVTAP 504

RESULT 2

P2C_PARTE
 ID_P2C_PARTE STANDARD; PRT; 300 AA.
 AC P49444;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Protein phosphatase 2C (EC 3.1.3.16) (PP2C).
 OS Parametrium tetraurelia.
 OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Penicillida;
 OC Parametrium.
 OX NCBI_TaxID=5888;
 RN [1]

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RX MEDLINE=95105156; PubMed=7806499;
 RA Klump S., Hanke C., Donella-Deana A., Beyer A., Kellner R.,
 RA Pina L.A., Schultz J.E.;
 RT "A membrane-bound protein phosphatase type 2C from Parametrium
 tetraurelia. Purification, characterization, and cloning.";
 RL J. Biol. Chem. 269:32774-32780(1994).

CC -!- FUNCTION: ENZYME WITH A BROAD SPECIFICITY.
 CC -!- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
 CC phosphate.
 CC -!- COFACTOR: BINDS 2 MAGNESIUM OR MANGANESE IONS (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Membrane-bound.
 CC -!- PTM: THE N-TERMINUS IS BLOCKED.
 CC -!- SIMILARITY: BELONGS TO THE PP2C FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: Z36985; CAB5448.1; -
 DR HSBP; P35813; IAG6.
 DR InterPro: IPR001932; PP2C-like.
 DR InterPro: IPR000222; PP2C.
 DR Pfam: PF00481; PP2C; 1.
 DR SMART: SM00331; PP2C StG; 1.
 DR SMART: SM00332; PP2C; 1.
 DR PROSITE: PS01032; PP2C; 1.
 KW Hydrolyase; Magnesium; Manganese; Membrane.
 FT METAL 37 37
 FT METAL 38 38
 FT METAL 57 57
 FT METAL 237 237
 FT METAL 289 289
 SQ SEQUENCE 300 AA; 33739 MW; BC0318B4F7724EC CRC64;

Query Match 7.4%; Score 190; DB 1; Length 300;
 Best Local Similarity 24.5%; Pred. No. 1e-05;
 Matches 66; Conservative 51; Mismatches 104; Indels 48; Gaps 9;

QY 64 LYGVNGYDGNRVTFVFAQRLSABELLGQLNAEHADVRVLLQAFDVERSFLESID 123
 DB 52 VFGVFDHGGRVAVQFVEKHFDVDELLKNK-----NFKEDQFEALKE 93
 QY 124 ALAEXASLSQSLPEGVPHQQLPROYOKILERLKTLEREIS-GGAAVAVAVLNNKLYVAN 182
 DB 94 TFLKMDL-----LTPEQKELNQYKAVDTESYAGCTANVALLIKNTLYVAN 142
 QY 183 VGNTRALLCKSTVDGLQVTLNDHTTENDELFRLSQGLDAGKIKOYGII CGQESTR 242
 DB 143 AGDSRSVLCRNNTN-----HDMVDHKPDNPEKSKITERAG---GVSDGRVAGNINLSRA 195


```

QY 243 IGDYKVGYYTDILLGSAKSKPIIAEPIHGAA--PLDGVGTGLVMSGLYKALEAAH 300
Db 196 LGDLEYXR-----DNKLRSEQLIIALPDVKKYTELTQD---KFLMGCGGVFETLNHOE 247

QY 301 GPGQANQEIA-AMIDTFEAKQTSIDAVAQ 328
Db 248 LLKQNSTIGQAQVTEELLKKAEDLLDQ 276

RESULT 3
P2C2 ARATH
ID P2C2 ARATH STANDARD; PRT; 399 AA.
AC P49598;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein phosphatase 2C (EC 3.1.3.16) (PP2C).
GN AT3G11410 OR F24K9.8.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=95116318; PubMed=7816619;
RA Kuromori T., Yamamoto M.;
RT "Cloning of cDNAs from Arabidopsis thaliana that encode putative
RT protein phosphatase 2C and a human Dr1-like protein by transformation
RT of a fission yeast mutant.";
RL Nucleic Acids Res. 22:5296-5301(1994).
RN [2]

SEQUENCE FROM N.A.
RP STRAIN=cv. Columbia;
RX MEDLINE=21016720; PubMed=11130713;
RA Salanoubat M., Lemcke K., Rieger M., Ansoerge W., Unsel M.,
RA Fartmann B., Valle G., Bloeker H., Perez-Alonso M., Obermaier B.,
RA Delseny M., Boutry M., Grivell L.A., Mache R., Puigdomenech P.,
RA De Simone V., Choigne N., Artiguenave F., Robert C., Brottier P.,
RA Wincker P., Cattolico L., Weissbach J., Saurin W., Quetier F.,
RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,
RA Wurmbach E., Drzonek H., Erfle H., Jordan N., Bangert S.,
RA Wiedemann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,
RA Vezzi A., D'Angelo M., Pallavicini A., Toppo S., Simonati B.,
RA Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordsiek G.,
RA Reichelt J., Scharfe M., Schoen O., Barges M., Terol J., Climent J.,
RA Navarro P., Collado C., Perez-Perez A., Ottenwaelder B., Duchemin D.,
RA Cooke R., Laudie M., Berger-Llauró C., Purnelle B., Masuy D.,
RA de Haan M., Maarse A.C., Alcaraz J.-P., Cortet A., Casacuberta E.,
RA Monfort A., Argirou A., Flores H., Rudd S., Zaccaria P., Mewes H.-W.,
RA Mannhaupt G., Haase D., Schoof H., Koo H.L., Tallon L.J., Jenkins J.,
RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
RA Rooney T., Rizzo M., Walts A., Uterback T., Fujii C.Y., Shea T.P.,
RA Cready T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,
RA Pal G., Millscher J., Sellers P., Gill J.B., Feldblyum T.V.,
RA Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,
RA Fraser C.M., Kaneo T., Nakamura Y., Sato S., Kato T., Asamizu E.,
RA Sasamoto S., Kimura T., Idesawa K., Kawashima K., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakayama S., Nakazaki N., Shino S., Takeuchi C., Wada T.,
RA Watanabe A., Yamada M., Yasuda M., Tabata S.;
RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis
RT thaliana.";
RN Nature 408:820-822(2000).
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate. BINDS 2 MAGNESIUM OR MANGANESE IONS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE PP2C FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D38109; BAA07287.1; -.
DR EMBL; AC008153; AAG51448.1; -.
DR HSSP; P35813; 1A6O.
DR InterPro; IPR001932; PP2C-like.
DR InterPro; IPR000222; PP2C.
DR Pfam; PF00481; PP2C; 1.
DR SMART; SM00331; PP2C SIG; 1.
DR SMART; SM00332; PP2C; 1.
DR PROSITE; PS01032; PP2C; 1.
KW Hydrolase; Magnesium; Manganese; Multigene family.
SQ SEQUENCE 399 AA; 43350 MW; 83B82E32FEC71D4D CRC64;

Query Match 6.8%; Score 175; DB 1; Length 399;
Best Local Similarity 25.5%; Pred. No. 0.00016;
Matches 93; Conservative 58; Mismatches 117; Indels 96; Gaps 20;

QY 7 SLQSEQPSWTDLPLCHLSGVGSASRSYSADGKGTES-HPPEDSWLKFRSENNCFLY 65
Db 91 SVTEAE--SFFSDVP-----KIGTTSVCGRRRDMEDAVSIHP---SFLQRNSENHHP-Y 138

QY 66 GVFNGYDGNRVTFVAQRLL-----SAELLQLQNAEADVRVLLQAPDVSFLE 119
Db 139 GYDFDGGHGSVHAEKCRERLHDIVKKEVEVNASD---EWTETWKRSGFQKMDKEVSORECNL 195

QY 120 SIDDAL-AEKASLSQLPEGVPHQLPQYQKILERLKLTEREISGGAMAVVAVLLNNKL 178
Db 196 VNGATRSKMKNSCRCELOS-----PQCDAV-----GSTAVSVVVTPEKI 234

QY 179 YVANVTNRALLCKSTVDGLQVTQLNVDHDTTENEDELRLSLQLGJDAKI-----KQVG 232
Db 235 IVSNCCDSRAVLCRNGV-----AIPLSVDHKPDRPDELRIRIQAG---GRVIYWDGARVLG 287

QY 233 IICQGESTERIGDYKVGYYTDILLGSAKSKPIIAEPIHGAAQPLDGVGTGLVMSGL 292
Db 288 VLA---MSRAIGDNVLIK-----PYVIPDPVTVTDRTD-EDECLIIASDGL 329

QY 293 YKAL--EAA-----HGPQANQEIAA-----MIDTEFAKQTSIDAVAQAVVD 332
Db 330 WDVVPNEATAGVARNCLRGAGAGDDSDAAHNACSDAALLTKLALAROSS-DNVSVVVVD 388

QY 333 RVKR 336
Db 389 LRKR 392

RESULT 4
P2C2 ARATH
ID P2C2 ARATH STANDARD; PRT; 423 AA.
AC O04719;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Protein phosphatase 2C AB12 (EC 3.1.3.16) (PP2C) (Abscisic acid-
DE insensitive 2).
GN AB12 OR AT5G57050 OR MHM17.19.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Landsberg erecta, and cv. Columbia;
RX MEDLINE=97308526; PubMed=9165752;
RA Leung J., Merlot S., Giraudat J.;
RT "The Arabidopsis ABSCISIC ACID-INSENSITIVE2 (AB12) and AB11 genes
RT encode homologous protein phosphatases 2C involved in abscisic acid

```


RA Schnabl S., Hiller R., Schmidt W., Lechharny A., Aubourg S.,
RA Cherdor F., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA Gibbons T., Weber N., Vandenbol M., Barges M., Terol J., Torres A.,
RA Peziz-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA Heijnen L., Schwarz S., Scholler P., Heber S., Francs P., Bielke C.,
RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
RA Farnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
RA Antonoli B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
RA Granat E., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
RA Chen E., Marra M., Martienssen R., McCombie W.R.;
RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
RT thaliana".
RN Nature 402:769-777(1999).
RN [5].
RP SEQUENCE FROM N.A.
RC STRAIN=ev. Columbia;
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the
RT SSP consortium (Salk/Stanford/PGSC)";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INVOLVED IN ABSICISIC ACID (ABA) SIGNALING PATHWAY.
CC -!- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC cofactor.
CC -!- COFACTOR: BINDS 2 MAGNESIUM OR MANGANESE IONS (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE PP2C FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
CC or send an email to license@sb-sib.ch).
CC -----
DR EMBL; X78886; CAA55484.1; -;
DR EMBL; U12856; AAA50237.1; -;
DR EMBL; X77116; CAA54383.1; -;
DR EMBL; AL049483; CAB39673.1; -;
DR EMBL; AL161564; CAB79463.1; -;
DR EMBL; AY035073; AAK59578.1; -;
DR HGSP; P35813; 1A6Q.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR001932; PP2C-like.
DR InterPro; IPR00222; PP2C.
DR Pfam; PF00481; PP2C; 1.
DR SMART; SM00331; PP2C_SIG; 1.
DR SMART; SM00332; PP2CC; 1.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS01032; PP2C; 1.
KW Hydrolase; Magnesium; Manganese; Multigene family; Calcium-binding.
FT CA BIND 93 104
FT DOMAIN 417 420
FT VARIANT 180 180
FT FT
FT FT
FT CONFLICT 24 24
FT CONFLICT 105 105
FT CONFLICT 1
SQ SEQUENCE 434 AA; 47505 MW; 4A4C54F04195F72 CRC64;

Query Match 6.5%; Score 167.5; DB 1; Length 434;
Best Local Similarity 21.8%; Pred.No. 0.00055;
Matches 69; Conservative 53; Mismatches 99; Indels 95; Gaps 11;

QY 55 KFRSENNCFYGVNGYDGNRVTNFVAQRISAELLGLQLNAAEHAAEDVRVVLQAFDWE 114

Db 163 RFPDQSAHFEGVYDGGSGQVANCYRMHLLA-----AEEI-AKEPMLCGDGTWLE 215
 Qy 115 RSFLSEIDDALEAKASQSQLPQGVPOQLPPQYQKTLERLKLTLEREISGNAVAVAVLL 174
 Db 216 KWKKALFNSFLRVDSIESVAPETV-----GSTSVVAVVF 250
 Qy 175 NKKLYVANVGTNRALLCKSTVDGLQVTLNVHDHTTNEDELFRLSQLGLDAGRIKQ-----230
 Db 251 PSHFIVANCNGDSRAVLCK-----GKTALPLSDVHKPDREDEAARIEAG---GKVIQWNGA 303
 Qy 231 --VGIICQSESTRIGDYKVKYGYTDDLLLSAAKSPRTIAEPIHGAQPLDGVGTGLVLM 288
 Db 304 RVFGVLA---MSRSIGDRIYK-----PSIIPDPDEVTAVKRV-EDDCLLILA 345
 Qy 289 SEGILYKALEAAHGPGQANQEI-----AAMIDTEFAKQT 321
 Db 346 SDGVMDVMTDEEACEMARKILLWHKQNAVAGDASLLADERRKEGKDPAAMSAAEYLSKL 405
 Qy 322 SL-----DAVAQAAYVD 332
 Db 406 AIQRGSKDNISVVVD 421
 RESULT 6
 ID P2C2 SCHPO STANDARD; PRT; 370 AA.
 AC Q09172;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Protein phosphatase 2C homolog 2 (EC 3.1.3.16) (PP2C-2).
 GN PTC2 OR SPCC1223.11.
 OS Schizosaccharomyces pombe (Fission yeast)
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 RN NCBI_TaxID=4896;
 RX [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Brooks J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Collins M., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Haidalo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy S., Niblett D., Odell C.,
 RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skellton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
 RA Weltegens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritzc C., Holzer C., Lehrach H., Reinhardt R., Pohl T.M.,
 RA Borzym K., Langer I., Beck A., Lehrach E., Wambutt R., Purnelle B.,
 RA Eger P., Zimmermann W., Wedler H., Mambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,

RA Shpakovski G.V., Usery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe";
 RL Nature 415:871-880(2002).
 CC - FUNCTION: HAS AN IMPORTANT ROLE IN OSMOTIC STABILITY AND CELL
 CC SHAPE CONTROL. IT MAY NEGATIVELY REGULATE THE OSMOSENSING SIGNAL
 CC TRANSMITTED THROUGH WIS1 MAP KINASE.
 CC - CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
 CC phosphate.
 CC - COFACTOR: BINDS 2 MAGNESIUM OR MANGANESE IONS (BY SIMILARITY).
 CC - SUBUNIT: MONOMER.
 CC - SIMILARITY: BELONGS TO THE PP2C FAMILY.
 CC - SIMILARITY: BELONGS TO THE PP2C FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: L34881; AAA67320.1; -
 CC EMBL: AL031579; CAA20880.1; -
 CC HSPB; P35813; IABQ. PP2C-like.
 CC InterPro: IPR001932; PP2C-like.
 CC InterPro: IPR00222; PP2C.
 CC Pfam: PF00481; PP2C; 1.
 CC SMART: SM00331; PP2C SIG; 1.
 CC SMART: SM00332; PP2C; 1.
 CC PROSITE: PS01032; PP2C; 1.
 CC HYDROLASE; Magnesium; Manganese; Multigene family.
 CC METAL 37 37 MANGANESE 1 (BY SIMILARITY).
 CC METAL 38 38 MANGANESE 1 (BY SIMILARITY).
 CC METAL 63 63 MANGANESE 1 AND 2 (BY SIMILARITY).
 CC METAL 233 233 MANGANESE 2 (BY SIMILARITY).
 CC METAL 282 282 MANGANESE 2 (BY SIMILARITY).
 CC SEQUENCE 370 AA; 40878 MW; DEDDB26F44AC50 CRC64;
 SO
 Query Match 6.2%; Score 160.5; DB 1; Length 370;
 Best Local Similarity 20.7%; Pred. No. 0.0013;
 Matches 69; Conservative 63; Mismatches 105; Indels 97; Gaps 16;
 QY 51 DSWLKF-----RSENNCF-----YGVNGYNGNVTNFVAQRL 84
 DB 19 DRMLHFGVSHQWGRISMEDAHCALNFTDSNSNPSTFSGVDFGHDGDAVAKYCRQHL 78
 QY 85 SAELLIGQLNHAADRVRLVLAFAVDERSFLESIDDALEKASLSQSPQEGVPHQL 144
 DB 79 -PDITKSPSPFKKNYD-----EALKSGFL-ADNMLMODRDMQ----- 115
 QY 145 PPQYKILERLKTLEREISGGAMAVAVLNNK-LYVANVGTNRALL-CKSTVDGLQVQ 202
 DB 116 -----EDPSGCTATTALIVDHQVIYCANAGDSRVLGRKGTAE-----P 154
 QY 203 LNVHTTENEDLFRLSQLG--LDAGKIKQYGIIGQESTRIGYKYKYVTTIDLLSA 260
 DB 155 LSFHDKPNNDVKARITAAAGFIDFGRVNA--GSLA---LSRAIGFEYKK-----DSSLP 204
 QY 261 ASKRPILAEPE--IHGAQPLDVGTFVLVMSGLYKALEAHAGPQAOEIAAMIDEFA 318
 DB 205 PEKQIVTAPEVDVINIDDD---EFLIACGIMDC-----KSSQVVEFVRGIV 253
 QY 319 KQTSIDAVAAQAVVDRVKIRHSDTPFASGGERAFQC 352
 DB 254 AROSLVEICENIMDRCIASNSSECGICDNMTIC 287
 RESULT 7
 ID PDP1 RAT STANDARD; PRT; 538 AA.
 AC O88483;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE [Pyruvate dehydrogenase [lipoamide]]-phosphatase 1, mitochondrial
 DE precursor (EC 3.1.3.43) (PDP 1) (Pyruvate dehydrogenase phosphatase,
 DE catalytic subunit 1) (PDP 1).
 GN PDP1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OK NCBI_Taxid=10116;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Heart;
 RX MEDLINE=98316337; PubMed=9651365;
 RA Huang B., Gudi R., Wu P., Harris R.A., Hamilton J., Popov K.M.;
 RT Isoenzymes of pyruvate dehydrogenase phosphatase: DNA-derived amino
 RT acid sequences, expression, and regulation.";
 RL J. Biol. Chem. 273:17680-17688(1998).
 CC - FUNCTION: CATALYZES THE DEPHOSPHORYLATION AND CONCOMITANT
 CC REACTIVATION OF THE ALPHA SUBUNIT OF THE E1 COMPONENT OF THE
 CC PYRUVATE DEHYDROGENASE COMPLEX.
 CC - CATALYTIC ACTIVITY: [Pyruvate dehydrogenase (lipoamide)] phosphate
 CC + H(2)O = [pyruvate dehydrogenase (lipoamide)] + phosphate.
 CC - COFACTOR: MAGNESIUM DEPENDENT AND CALCIUM-STIMULATED.
 CC - SUBUNIT: HETERODIMER OF A CATALYTIC SUBUNIT AND A P4D PROTEIN OF
 CC UNKNOWN FUNCTION (BY SIMILARITY).
 CC - SUBCELLULAR LOCATION: Mitochondrial matrix.
 CC - TISSUE SPECIFICITY: HIGHLY EXPRESSED IN SKELETAL MUSCLE.
 CC - SIMILARITY: BELONGS TO THE PP2C FAMILY.
 CC - SIMILARITY: BELONGS TO THE PP2C FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AF062740; AAC40167.1; -
 CC InterPro: IPR001932; PP2C-like.
 CC InterPro: IPR00222; PP2C.
 CC Pfam: PF00481; PP2C; 1.
 CC SMART: SM00331; PP2C SIG; 1.
 CC SMART: SM00332; PP2C; 1.
 CC PROSITE: PS01032; PP2C; 1.
 CC HYDROLASE; Mitochondrion; Transit peptide; Magnesium; Calcium-binding.
 CC TRANSIT 1 71 MITOCHONDRION (BY SIMILARITY).
 CC CHAIN 72 538 [PYRUVATE DEHYDROGENASE [LIPAMIDE]]-
 CC PHOSPHATASE 1.
 CC SEQUENCE 538 AA; 61207 MW; D546096229B73564 CRC64;
 SO
 Query Match 6.0%; Score 154; DB 1; Length 538;
 Best Local Similarity 24.7%; Pred. No. 0.0067; Indels 140; Gaps 17;
 Matches 89; Conservative 38; Mismatches 93; Indels 140; Gaps 17;
 QY 40 DGKGTES-----HPEDSWLKFRSENNC-----FLYGVNGYNGNVTNFVAQ 82
 DB 102 DGNVSSILGFDNSRNPANAPIEDR---RSATTCIGTQGMILGVFDGAGACGQAVSE 157
 QY 83 R-----LSAELLIGQLNA----- 95
 DB 158 RLRYIAVSLPPELTLELENVESGRALLPILQWHPNDYFSKASKLYENGLRTYWQ 217
 QY 96 -----EHAEDVRRVLLQAFDVERSFLESIDDALEKASLSQSPQEGVPHQLPQ 147
 DB 218 ELIDNTGESADIDVKEALINAF-----KRLDNDI---SLEAQVSD-----PNS 258
 QY 148 YOKILERLKTLEREISGGAMAVAVLNNKLYVANVGTNRALLCKSTVDG-LQVQLNVND 206
 DB 259 FLNYL-----VLRAFSGATACVAHVDGVDHVAANTGDSRAVLGVQEDGSSAVTILSD 313
 QY 207 HTTENEDLFRLSQLDAGK-----IKQVGIIGQESTRIGYKYKYVTTIDLLS- 259
 DB 314 HNAQNERELQRLK---LEHPKNEAKSVVQDRLILGLMFPFAFGVDFKFM---SIDLQKR 367

```
QY 260 AAKSKP-----IIAPEI--HGAOPLDGVTFGLVMSGLYKAL 296
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 368 VIESGPDQNDNEYTKFPNHYTPPVLTAEVTVYHRLRPQD---KFLVLATDGLWETM 424

RESULT 8
ID P2C1_CAEEL STANDARD; PRT; 491 AA.
AC P49595;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Probable protein phosphatase 2C F42G9.1 (EC 3.1.3.16) (PP2C).
GN F42G9.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Taich A., Waterston R.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
CC -!- COFACTOR: BINDS 2 MAGNESIUM OR MANGANESE IONS (BY SIMILARITY).
CC -!- COFACTOR: BINDS 2 MAGNESIUM OR MANGANESE IONS (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE PP2C FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U00051; AAA91358.1; -.
CC HSSP; P35813; 1A6Q.
CC WormPep; F42G9.1; CE07231.
CC InterPro; IPR001932; PP2C-like.
CC InterPro; IPR000222; PP2C.
CC Pfam; PF00481; PF2C; 1.
CC SMART; SM00331; PP2C_SIG; 1.
CC SMART; SM00332; PP2CC; 1.
CC PROSITE; PS01032; PP2C; 1.
CC KW Hypothetical protein; Hydrolase; Magnesium; Manganese.
FT METAL 37 37 MANGANESE 1 (BY SIMILARITY).
FT METAL 38 38 MANGANESE 1 (BY SIMILARITY).
FT METAL 57 57 MANGANESE 1 AND 2 (BY SIMILARITY).
FT METAL 428 428 MANGANESE 2 (BY SIMILARITY).
FT METAL 477 477 MANGANESE 2 (BY SIMILARITY).
SQ SEQUENCE 491 AA; 53141 MW; F110D12E343953D6 CRC64;

Query Match 5.9%; Score 151.5; DB 1; Length 491;
Best Local Similarity 21.3%; Pred. No. 0.0077;
Matches 77; Conservative 63; Mismatches 118; Indels 103; Gaps 13;

QY 5 RRSLLQSQSQSWTDDPLCHLSGVGSASRNSYADGKGTESHPPEDSWLKFRSENNCFL 64
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 229 QKSPIQSEAKSKSE-----TDAETAPSSSSGGVGVATSEEDDSDFVADE---- 277
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:

QY 65 GYVFNGYDGNRVNVAQLSAELLGLQNAHAEADVRRVLLQAFDVVRSFLESIDDA 124
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 278 -----BEDDEDAED-----EQSDDEMVDGS 297

QY 125 LAEKASLOSQLEPGVPQHPQVQKILRLKTLREITSGGAMAVVLLNNKLYVANVG 184
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 298 LA-----PLLLGGGAEEVGE-----DSGTTACVCLVGKDKVIVANAG 335

QY 195 TNRAALLCKSTVDGLQVLTQNLVDHTTENEDELFRLSQLGLDAGKIKQVGIICQESTRIG 244
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 336 DSRAVLCRNG----KAVDLSVDHKEPEDEVETNRHAAG---GQIEDGRVNGGLNSRAFG 388
```

```
QY 245 DYKVKYGYTDILLSSAASKPIIABPEIHGAOPLDGVTFGLVMSGLYKALEAAHGPQ 304
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 389 DYAKKN-QELGL-----KEQMITALPDVK-IEALTPEDEFIVVACDGIWNSME----- 435

QY 305 ANOEIAAMIDTEFAKOTSILDAVAQAQVDRVKRIHSDHTFASGERARFCPRHEDMTLLVRN 364
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 436 -SQQVDFVRDLLAKGSSCAEVCDALCDACLDADSTDGGGTG-----C---DNMTVICTT 485

QY 365 F 365
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 486 F 486

RESULT 9
ID P2C2_CAEEL STANDARD; PRT; 356 AA.
AC P49596;
DT 01-FEB-1996 (Rel. 33, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable protein phosphatase 2C T23F11.1 (EC 3.1.3.16) (PP2C).
GN T23F11.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Swinburne J.;
RL Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS.
RA Durbin R.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -!- COFACTOR: BINDS 2 MAGNESIUM OR MANGANESE IONS (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE PP2C FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z46343; CAA86456.2; -.
CC HSSP; P35813; 1A6Q.
CC WormPep; T23F11.1; CE24009.
CC InterPro; IPR001932; PP2C-like.
CC InterPro; IPR000222; PP2C.
CC Pfam; PF00481; PF2C; 1.
CC SMART; SM00331; PP2C_SIG; 1.
CC SMART; SM00332; PP2CC; 1.
CC PROSITE; PS01032; PP2C; 1.
CC KW Hypothetical protein; Hydrolase; Magnesium; Manganese.
FT METAL 37 37 MANGANESE 1 (BY SIMILARITY).
FT METAL 38 38 MANGANESE 1 (BY SIMILARITY).
FT METAL 59 59 MANGANESE 1 AND 2 (BY SIMILARITY).
FT METAL 228 228 MANGANESE 2 (BY SIMILARITY).
FT METAL 277 277 MANGANESE 2 (BY SIMILARITY).
SQ SEQUENCE 356 AA; 39064 MW; EDCB0841CFB026B5 CRC64;

Query Match 5.8%; Score 148.5; DB 1; Length 356;
Best Local Similarity 20.4%; Pred. No. 0.0076;
Matches 64; Conservative 63; Mismatches 105; Indels 81; Gaps 14;

QY 29 VGSASRNSYSADGKGTESHPPEDSWLKFRSENNCFLYGVFNGYDGNRVNVAQLRSABL 88
      :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 24 VGSNCMQGVGRVDMEDAHTH-----LLSLPDDPKCAFFAYVDYDGGGSKVSQYSGINLHKV 78
```

```

QY 89 LIGQLNHAHADVRVRLIQAFDVERSFLESTIDDLAKSAQSLQPEGVPHQLPPQY 148
DB 79 V---AOKESSEGKMK-----EALKEGPLE-----LDQOM-----104
QY 149 QKILRLKLEIEISGAMAVVAVLNNKLYVANVTNRALLCKSTVDGLQVTQLNVDT 208
DB 105 -----RVDEETADVSGTAAVVVLLIKESGVYCGNAGDSRAV---SSVVG-EARPLSPDHK 155
QY 209 TENEDLFLPSQLGDAGKIKOVGIIICGQES-TRRIGDYKVKYGTDDIDLSAAKSPTI 267
DB 156 PSHEFEARRI-----IAAGWVEFNRRNGMIALSRALGDFAFK-----NCDTKP--199
QY 268 ASEPETHGAPLDGVNG-----FLVMSEGLYKALEAHGPGANOFIAMIDTEPAKQ 320
DB 200 AAEQIVTAF-P-DVITDKLTPDHEFVLACGIWDVW-----TNQEVDPFVREKLAEK 250
QY 321 TSLDAVQAQVDR 333
DB 251 RDPQSIQCELLTR 263

RESULT 10
PDP2_HUMAN STANDARD; PRT; 529 AA.
ID PDP2_HUMAN
AC Q9P2J9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE [Pyruvate dehydrogenase (lipoamide)]-phosphatase 2, mitochondrial precursor (EC 3.1.3.43) (PDP 2) (Pyruvate dehydrogenase phosphatase, catalytic subunit 2) (PDP 2).
GN PDP2 OR KIAA1348
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20181126; PubMed=10718198;
RA Nagase T., Kikuno R., Ishikawa K.-I., Hirose M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XVI. The complete sequences of 150 new cDNA clones from brain which code for large proteins in vitro."
RL DNA Res. 7:65-73(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CATALYZES THE DEPHOSPHORYLATION AND CONCOMITANT REACTIVATION OF THE ALPHA SUBUNIT OF THE E1 COMPONENT OF THE PYRUVATE DEHYDROGENASE COMPLEX (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: [Pyruvate dehydrogenase (lipoamide)] phosphate + H(2)O = [pyruvate dehydrogenase (lipoamide)] + phosphate.
CC -!- COFACTOR: MAGNESIUM DEPENDENT (BY SIMILARITY).
CC -!- SUBUNIT: HETERODIMER OF A CATALYTIC SUBUNIT AND A P2D PROTEIN OF UNKNOWN FUNCTION (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Mitochondrial matrix (By similarity).
CC -!- SIMILARITY: BELONGS TO THE P2C FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation- the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb.ch/announce/ or send an email to license@isb.ch).
CC -----
CC EMBL; AB037769; BAA92586.1; ALT_INIT.
CC EMBL; BC028030; AAA82030.1; -.
CC InterPro; IPR001932; P2C-like.
CC InterPro; IPR000222; P2C.

```

```

DR Pfam; PF00481; P2C; 2.
DR SMART; SM00331; P2C_SIG; 1.
DR SMART; SM00332; P2C; 1.
DR PROSITE; PS01032; P2C; 1.
FT K1 Hydroxylase; Mitochondrion; Transit peptide; Magnesium.
FT TRANSIT 1 66 MITOCHONDRION (POTENTIAL).
FT CHAIN 67 529 [PYRUVATE DEHYDROGENASE (LIPOAMIDE)] -
FT SEQUENCE 529 AA; 59978 MW; 252CAEBDAF61A5C CRC64;
SQ
Query Match 5.7%; Score 147; DB 1; Length 529;
Best Local Similarity 22.8%; Pred. No. 0.017;
Matches 84; Conservative 50; Mismatches 96; Indels 138; Gaps 17;
QY 45 ESHPEPDSWLKFRS-----ENNCFYGVNGYDGNRVTFVAAQL- 84
DB 99 ESRVP-NSVLRFSNOLAANSFVEDRRGVASCLOTGMLFGIFDGHGACAAQAVSERLF 157
QY 85 -SAILLIGQLNHAHA-----DYRVLLQAFDVERSF 118
DB 158 YVAVSLMSHQTLHEMGANESKPLPILHMLKHPGDSIYKDVSVHLDRVYWQELL 217
QY 119 E-----SIDALA-----EKASLQSLPEGVPHQLPPQYQKILRLKLEIEIS- 163
DB 218 DLMEMGLSIEELNMSFQRLDSDISLEIQAP-----LEDEVTR 256
QY 164 -----GGMAVAVLNNKLYVANVTNRALLCKSTVDGL-QVTQLNVDTTENED 214
DB 257 NLSLQVAFSGATCMAMHVDGILHVNAGDCRAILGVQEDNGWMSCLPLTRDHMANNQAE 316
QY 215 LFLPSQLGIDA-----GKIKOVGIIICGQESTRIGDYKVKY-----GYTDI 255
DB 317 LSRKRHPESDRTTIMEDRLGVLP-PC---RPFQVQLKSKELQSLIERGF-NT 370
QY 256 DLISAASKRP-----IIAEPFI--HGAQPLDVGTVGLVMSSEGLYALBAHGPQA 305
DB 371 EALNIQGFPPHYTPPYLTAEPVYHRLRPQD---KLVLASDLMML-----S 419
QY 306 NQETIAMT 313
DB 420 NEDVRLV 427

RESULT 11
PDP1_BOVIN STANDARD; PRT; 538 AA.
ID PDP1_BOVIN
AC P35816;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE [Pyruvate dehydrogenase (lipoamide)]-phosphatase 1, mitochondrial precursor (EC 3.1.3.43) (PDP 1) (Pyruvate dehydrogenase phosphatase, catalytic subunit 1) (PDP 1).
GN PDP1 OR PDP.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bos.
OX NCBI_TaxId=9913;
RN [1]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RX MEDLINE=93378941; PubMed=8396421;
RA Lawson J.E., Niu X.-D., Browning K.S., Trong H.L., Yan J., Reed L.J.;
RT "Molecular cloning and expression of the catalytic subunit of bovine pyruvate dehydrogenase phosphatase and sequence similarity with RT protein phosphatase 2C."
RL Biochemistry 32:8987-8993(1993).
CC -!- FUNCTION: CATALYZES THE DEPHOSPHORYLATION AND CONCOMITANT REACTIVATION OF THE ALPHA SUBUNIT OF THE E1 COMPONENT OF THE PYRUVATE DEHYDROGENASE COMPLEX.
CC -!- CATALYTIC ACTIVITY: [Pyruvate dehydrogenase (lipoamide)] phosphate + H(2)O = [pyruvate dehydrogenase (lipoamide)] + phosphate.
CC -!- COFACTOR: MAGNESIUM DEPENDENT AND CALCIUM-STIMULATED.

```

```
CC -!- SUBUNIT: HETERODIMER OF A CATALYTIC SUBUNIT AND A PAD PROTEIN OF
CC UNKNOWN FUNCTION.
CC -!- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -!- SIMILARITY: BELONGS TO THE PP2C FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L18966; AAA30697.1; ALT_INIT.
CC PIR; A48692; A48692.
CC DR InterPro; IPR001932; PP2C-like.
CC DR InterPro; IPR000222; PP2C.
CC DR Pfam; PF00481; PP2C; 1.
CC DR SMART; SM00331; PP2C_Sig; 1.
CC DR SMART; SM00332; PP2CG; 1.
CC DR PROSITE; PS01032; PP2C; 1.
CC KW Hydrolase; Mitochondrion; Transit peptide; Magnesium; Calcium-binding.
CC FT TRANSIT 1 71 MITOCHONDRION.
CC FT CHAIN 72 538 [PYRUVATE DEHYDROGENASE (LIPOAMIDE)]-
CC FT PHOSPHATASE 1.
CC SQ SEQUENCE 538 AA; 61184 MW; 1A1C219AD8C3DAE3 CRC64;
CC -----
CC Query Match 5.6%; Score 145; DB 1; Length 538;
CC Best Local Similarity 24.2%; Pred. No. 0.024;
CC Matches 87; Conservative 39; Mismatches 94; Indels 140; Gaps 17;
CC -----
CC QY 40 DGKTES-----HPPDSWLKFRSENNC-----FLYFNGYDGNRTNPFVAQ 82
CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
CC Db 102 DGNVSSVLGFDNSQLPANAPIEDR---RSAATCLQTRGMILGVFGHAGCACSQAVSE 157
CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
CC QY 83 R-----LSAELLIGQLNA-----
CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
CC Db 158 RLFFVYVSLPHTLLEIENAVESGRALLPQLQWKHPNDYFSKASLYFNLSRTYMQ 217
CC -----
CC QY 96 -----EHAEDVRRVLLOAFDVVERSFLSDIDLAELAKSLOSLPGVQHQHLPQ 147
CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
CC Db 218 ELIDNTGESTDIDVKEALINAF-----KRLDNDI-----SLEAQVGD-----PNS 258
CC -----
CC QY 148 YOKILERLKTREISGAMAVAVLLNNKLYVANVTNRALLCKSTVDG-LQVTLQNVLD 206
CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
CC Db 259 FLNVL-----VLRVAFSGATACVAHVDVCDLVHANTGDSRAMLGQVEDGSWAVTLNSD 313
CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
CC QY 207 HTTENEDFLRLSQLGLDAGK-----IKQVGIICGOESTRRIGDYKVYGYTDIDLLS- 259
CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
CC Db 314 HNAQNEREVERLK---LHPRKNEAKSVVKQDRLLGLLMPFRAFGDVKFKW---SIDLQKR 367
CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
CC QY 260 AAKSKP-----IIARPEI--HCAQPLDGVTFVLVLMSEGLYKAL 296
CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
CC Db 368 VIESGFDQLNDNEYTKFPPNYTPPYLTABPEVTHRLRPQD---KFLVLTADGLWETM 424
CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
CC -----
CC RESULT 12
CC P2C3_SCHPO STANDARD; PRT; 414 AA.
CC AC Q09173;
CC DT 01-NOV-1995 (Rel. 32, Created)
CC DT 15-JUN-2002 (Rel. 41, Last sequence update)
CC DE Protein phosphatase 2C homolog 3 (EC 3.1.3.16) (PP2C-3).
CC GS PTC3 OR SPAC2G11.07C.
CC ON Schizosaccharomyces pombe (Fission yeast).
CC OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
CC OC Schizosaccharomycetales; Schizosaccharomycetaceae;
CC OC Schizosaccharomycetes.
CC OK NCBI_TaxID=4896;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=972;
```

```
RX MEDLINE=95163582; PubMed=7859738;
RA Shiozaki K., Russell P.;
RT "Counteractive roles of protein phosphatase 2C (PP2C) and a MAP
RT kinase kinase homolog in the osmoregulation of fission yeast.";
RL EMBL J. 14:492-502(1995).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21849401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders R., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -!- FUNCTION: HAS AN IMPORTANT ROLE IN OSMOTIC STABILITY AND CELL
CC SHAPE CONTROL. IT MAY NEGATIVELY REGULATE THE OSMOSENSING SIGNAL
CC TRANSMITTED THROUGH WIS1 MAP KINASE.
CC -!- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -!- COFACTOR: BINDS 2 MAGNESIUM OR MANGANESE IONS (BY SIMILARITY).
CC -!- SUBUNIT: MONOMER.
CC -!- SIMILARITY: BELONGS TO THE PP2C FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L34882; AAA67321.1; -.
CC DR EMBL; Z54354; CAA91172.1; -.
CC DR HSSP; P35813; IA6Q.
CC DR InterPro; IPR001932; PP2C-like.
CC DR InterPro; IPR000222; PP2C.
CC DR Pfam; PF00481; PP2C; 1.
CC DR SMART; SM00331; PP2C_Sig; 1.
CC DR SMART; SM00332; PP2CG; 1.
CC DR PROSITE; PS01032; PP2C; 1.
CC KW Hydrolase; Magnesium; Manganese; Multigene family.
CC FT METAL 37 37 MANGANESE 1 (BY SIMILARITY).
CC FT METAL 38 38 MANGANESE 1 (BY SIMILARITY).
CC FT METAL 62 62 MANGANESE 1 AND 2 (BY SIMILARITY).
CC FT METAL 230 230 MANGANESE 2 (BY SIMILARITY).
CC FT METAL 279 279 MANGANESE 2 (BY SIMILARITY).
CC FT CONFLICT 196 196 K -> T (IN REF. 1).
CC SQ SEQUENCE 414 AA; 44856 MW; EFP3A416625A2B11 CRC64;
CC -----
CC Query Match 5.6%; Score 143.5; DB 1; Length 414;
CC Best Local Similarity 21.9%; Pred. No. 0.02;
CC Matches 62; Conservative 49; Mismatches 105; Indels 67; Gaps 12;
CC STRAIN=972;
```

```

QY 65 XGVFNGYDGNRYTNFYAQRSLSLGOLNABEADRVRLLOAFDVERSFILESIDA 124
DB 58 FAVYVGHGDGDKVAKWCGSNLPQIL---EKNPDPFGKDF-----VNLKSSFLN----- 102
QY 125 LAEKASLSQSLPEGVPOHQLPPQYOKILERLKTLEISGAMAVAVLNNKLYVANVG 184
DB 103 -ADKALDD-----QFITDP-----SGCTATVVLRVGNKLYCANAG 138
QY 185 TNRALLCKSTVDGLQVTLQVNDHTTENDELFRLSQLG--LDAGKIKOVGIIICQGESTR 242
DB 139 DSRTVLGSKGI---AKPLSADHKPSNEAEKARICAGGFVDFGRVNG-----NLALSRA 189
QY 243 IDDYKVKYVYTDILLSAKSKPFIAPRPHGAQPLDGTGFLVMSGGLYKALEAAGP 302
DB 190 IDDFEFKSNLE-----PEKQIVTALPQVVEHETTD--DEFVVLACQGIWDC----- 235
QY 303 GOANOEIAMIIDTEFAKQTSLDVAQAADVVRKRIHSDTFASG 345
DB 236 -KTSQOVIIEFVRGIVAGTSLERKIAENLMDNC--IASDIETTG 275

RESULT 13
PDP1_HUMAN STANDARD; PRT; 538 AA.
ID PP1_HUMAN
AC Q9P0I1
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE [Pyruvate dehydrogenase (lipoamide)]-phosphatase 1, mitochondrial
DE precursor (EC 3.1.3.43) (PDP 1) (Pyruvate dehydrogenase phosphatase,
DE catalytic subunit 1) (PDP1C 1).
GN PDP1 OR PDP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP TISSUE=Adrenal gland;
RC MEDLINE=20402571; PubMed=10931946;
RA Hu R.-M., Han Z.-G., Song H.-D., Peng Y.-D., Huang Q.-H., Ren S.-X.,
RA Gu B.-W., Dai M., Mao Y.-F., Gao G.-F., Rong R., Ye M., Zhou J.,
RA Xu S.-H., Gu J., Shi J.-X., Jin W.-R., Zhang C.-K., Wu T.-M.,
RA Huang G.-Y., Chen Z., Chen M.-D., Chen J.-L.;
RT "Gene expression profiling in the human hypothalamus-pituitary-adrenal
RT axis and full-length cDNA cloning.";
RT Proc. Natl. Acad. Sci. U.S.A. 97:9543-9548(2000).
RL
CC - FUNCTION: CATALYZES THE DEPHOSPHORYLATION AND CONCOMITANT
CC REACTIVATION OF THE ALPHA SUBUNIT OF THE EI COMPONENT OF THE
CC PYRUVATE DEHYDROGENASE COMPLEX (BY SIMILARITY).
CC - CATALYTIC ACTIVITY: [Pyruvate dehydrogenase (lipoamide)] phosphate
CC + H(2O) = [pyruvate dehydrogenase (lipoamide)] + phosphate.
CC - COFACTOR: MAGNESIUM DEPENDENT AND CALCIUM-STIMULATED (BY
CC SIMILARITY).
CC - SUBUNIT: HETERODIMER OF A CATALYTIC SUBUNIT AND A P2D PROTEIN OF
CC UNKNOWN FUNCTION (BY SIMILARITY).
CC - SUBCELLULAR LOCATION: Mitochondrial matrix (By similarity).
CC - SIMILARITY: BELONGS TO THE PP2C FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC
CC EMBL; AF155661; AAF67480.1; ALT_INIT.
CC InterPro; IPR001932; PP2C-1like.
CC InterPro; IPR000222; PP2C.
CC Pfam; PF00481; PP2C; 1.

```

```

DR SMART; SM00331; PP2C_SIG; 1.
DR SMART; SM00332; PP2C; 1.
DR PROSITE; PS01032; PP2C; 1.
KW Hydroxylase; Mitochondrion; Transit peptide; Magnesium; Calcium-binding.
FT TRANSIT 1 71 MITOCHONDRION (BY SIMILARITY).
FT CHAIN 72 538 [PYRUVATE DEHYDROGENASE (LIPOAMIDE)]-
FT PHOSPHATASE 1.
SQ SEQUENCE 538 AA; 61219 MW; 1A3E512472540432 CRC64;

Query Match 5.5%; Score 143; DB 1; Length 538;
Best Local Similarity 24.7%; Pred. No. 0.032;
Matches 84; Conservative 38; Mismatches 90; Indels 128; Gaps 16;

QY 48 PPDSDMKFRSENNC-----FLYGVNNGYDGNRYTNFYAQRSL-----SAELLGOL 93
DB 122 P1BDR---RSAATCLOTRGMILGVDFGHAGCAMSQAVSERLFYYIAGSLVPHETLLEIE 177
QY 94 NA-----EHAEDRVRLV 107
DB 178 NAVESGRALLPIQWKKHPNDYFSKASLYFNSLRTYQELIDLTNGESTIDVKEALI 237
QY 108 QAFDVERSFLESIDDALEKASLSQSLPEGVPOHQLPPQYOKILERLKTLEISGAM 167
DB 238 NAF-----KRLNDI-----SLBAQVGD-----FNSFLNYL-----VLRAFSQAT 273
QY 168 AVAVAVLNNKLYVANVTNRALLCKSTVDG-LQVTLQVNDHTTENDELFRLSQGLDG 226
DB 274 ACVAHVDGVDLHVANTGDSRAMLVQGEEDGSWSAVTLSDNHAQNERELERLK--LEHP 330
QY 227 K-----IKOVGIIICQGESTRRIGDYKVKVGYTDDILLS-AAKSKP----- 265
DB 331 KSEAKSVKQDRLLGILMPFRAGDYKFKW---SIDLQKRVIESGPDQLNDNEYTKFIP 387
QY 266 -----IIAEPI--HGAQPLDGTGFLVMSGLYKAL 296
DB 388 NYHTPFLTAPEPVTYHRLRPQD---KFLVLTADGLMETM 424

RESULT 14
P2C3_ARATH STANDARD; PRT; 388 AA.
ID P2C3_ARATH
AC P49599; O9M0J6; Q9STP6;
DT 01-FEB-1996 (Rel. 33, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein phosphatase 2C PPH1 (EC 3.1.3.16) (PP2C).
GN PPH1 OR AT4G27800 OR T27E11.40.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA MEDLINE=96026352; PubMed=7569999;
RA Schena M., Shalon D., Davis R.W., Brown P.O.;
RT "Quantitative monitoring of gene expression patterns with a
RT complementary DNA microarray.";
RL Science 270:467-470(1995).
RP
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA MEDLINE=20083488; PubMed=10617198;
RA Mayer K.F.X., Scheller C., Wambutt R., Ertlan K.-D., Terry N.,
RA Pohl T., Duesterhoef A., Stiekema M., Ertlan K.-D., Terry N.,
RA Harris B., Anesorge W., Brandt P., Grivell L.A., Rieger M.,
RA Weichselgartner M., de Simone V., Obermaier B., Weiche R., Mueller M.,
RA Kreis M., Deleney M., Putzdomenech P., Watson M., Schmidheini T.,
RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA Vos P., Hebeisel J., Zimmermann W., Wedler H., Ridley P.,
RA Langham S.-A., McCullagh B., Bihlham L., Robben J.,
RA van der Schueren J., Grymoprez B., Chuang Y.-J., Vandenbussche F.,

```


RA Braeken M., Weljens I., Voet M., Bastiaens I., Aert R., Defoor E.,
 RA Weitzneger T., Bothe G., Ramsperger U., Hilbert H., Braun M.,
 RA Holzer E., Brandt A., Peters S., van Staveren M., Dirks W.,
 RA Moeljan P., Klein Lankhorst R., Rose M., Hauf J., Koeter P.,
 RA Berner S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
 RA De Keyser A., Buyschaert C., Gielen J., Villaroel R., De Clercq R.,
 RA Van Montegu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
 RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McLay K., Mayes R.,
 RA Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann T.,
 RA Borkova D., Blocker H., Scharfe M., Grimm M., Loehnert T.-H.,
 RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,
 RA Neumann S., Argirion A., Vitale D., Liguori R., Piravandi E.,
 RA Massenot O., Quigley F., Clabaud G., Muendlein A., Feiber R.,
 RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
 RA Chedfor F., Cooke R., Berger C., Monfort A., Casacuberta E.,
 RA Gibbons T., Weber N., Vandenbol M., Bagues M., Terol J., Torres A.,
 RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacón D., Jesse T.,
 RA Heijnen L., Schwarz S., Scholler P., Heber S., Frances P., Biele C.,
 RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
 RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
 RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
 RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
 RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
 RA Latrelle P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
 RA Manx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
 RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
 RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
 RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
 RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
 RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
 RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
 RA Granat S., Shohdy N., Hasegawa A., Hameed A., Lochi M., Johnson A.,
 RA Chen E., Marra M., Martienssen R., McCombie W.R.,
 RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
 RT thaliana";
 RL Nature 402:769-777(1999).
 CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
 CC phosphate.
 CC -1- COFACTOR: BINDS 2 MAGNESIUM OR MANGANESE IONS (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE PP2C FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; U34803; AAA92889.1; -;
 CC DR EMBL; AL161571; CAB81429.1; -;
 CC DR EMBL; AL078579; CAB43968.1; ALT_SEQ.
 CC DR InterPro; IPR001932; PP2C-like.
 CC DR InterPro; IPR000222; PP2C.
 CC DR Pfam; PF00481; PP2C; 1.
 CC DR SMART; SM00331; PP2C_Sig; 1.
 CC DR SMART; SM00332; PP2C; 1.
 CC DR PROSITE; PS01032; PP2C; 1.
 CC DR Hydrolase; Magnesium; Manganese; Multigene family.
 CC KW METAL 93 93 MANGANESE 1 AND 2 (BY SIMILARITY).
 CC FT METAL 296 296 MANGANESE 2 (BY SIMILARITY).
 CC FT METAL 326 326 MANGANESE 2 (BY SIMILARITY).
 CC FT CONFLICT 1 4 MALL -> SNSS (IN REF. 2).
 CC FT CONFLICT 69 70 RD -> QY (IN REF. 2).
 CC FT CONFLICT 271 271 D -> N (IN REF. 2).
 CC FT CONFLICT 353 353 W -> C (IN REF. 2).
 CC FT CONFLICT 360 388 RQNVVELVQAATTGLVTVGIMWSHLS -> TKRSQ
 CC (IN REF. 2).
 CC FT SEQUENCE 388 AA; 42719 MW; 496ECC786AEB802 CRC64;
 CC
 CC Query Match 5.4%; Score 139.5; DB 1; Length 388;
 CC Best Local Similarity 21.3%; Pred. No. 0.034;

Matches 64; Conservative 58; Mismatches 99; Indels 79; Gaps 13;
 QY 63 ELYGVGVNGYDGNRVTFNFAOQLRSAB-----LLLGQINAEHAEADVRVLLQAFDVVERSF 117
 DB 86 FSYAAVFDGHAGSSSVKFLREELYKECVGALQAGSLNGGDFAAIKALIKAFESVDRNL 145
 QY 118 LESIDDAEAKASLQSLPQGVQPQOLPQYQKILERLKTIERETSGGAMAVVALLNKK 177
 DB 146 LKWLK-----ANGDEDESGSTATVMIIRDV 172
 QY 178 LVVANVTNRALLCKSTVDGLQVLTQNDVHTTENE-----DELFRLSQLGLDAGKIKQVG 232
 DB 173 SFTAHIGDSCAVLSRSG---QIEBLTDVHRPYGSSRAAIQEVKRVKEAG---GWIVN-G 224
 QY 233 IICQGES-TRRTGDKVKVGYTDI-----DILLSA--KSKPIIAEPETHGAQP 277
 DB 225 RICGDIASVRAFGRDIFRTKKNMILKKGVDGWRSEKVFVSRIEFGDMVATPDIFQV-P 283
 QY 278 LDGVTGFLVMEGLYKALEAAHGPQANQEAAMIDTEFAK-----QTSLDAVAQAVVDR 333
 DB 284 LTSDFEIIILASDGLWDMYKSS-----DWSVVRDLRKHGNVQLACESLAQVALDR 335
 RESULT 15
 P2C4 SCHPO STANDARD; PRT; 383 AA.
 ID P2C4 SCHPO QSURO2;
 AC 014156; QSURO2;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Protein phosphatase 2C homolog 4 (EC 3.1.3.16) (PP2C-4).
 GN PTC4 OR SPAC4A8.03C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A., SUBUNIT, SUBCELLULAR LOCATION, AND FUNCTION.
 RC STRAIN=972;
 RX MEDLINE=99365157; PubMed=10436019;
 RA Gaits F., Russell P.;
 RT "Vacuole fusion regulated by protein phosphatase 2C in fission
 RL [2].
 RL Mol. Biol. Cell 10:2647-2654 (1999).
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grynoprte B.,
 RA Weltjens I., Vansteels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehrach H., Wambutt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Motier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;

```

RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -1- FUNCTION: HAS A ROLE IN THE REGULATION OF VACUOLE FUSION.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- COFACTOR: BINDS 2 MAGNESIUM OR MANGANESE IONS (BY SIMILARITY).
CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, VACUOLAR.
CC -1- SIMILARITY: BELONGS TO THE PP2C FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF140285; AAD27651.1; -.
DR EMBL; Z98762; CAB58554.1; -.
DR InterPro; IPR001932; PP2C-like.
DR InterPro; IPR000222; PP2C.
DR Pfam; PF00481; PP2C; 1.
DR SMART; SM00331; PP2C_SIG; 1.
DR SMART; SM00332; PP2C; 1.
DR PROSITE; PS01032; PP2C; 1.
DR HydroLase; Magnesium; Manganese; Membrane; Multigene family.
KW METAL 92 92 MANGANESE 1 AND 2 (BY SIMILARITY).
FT METAL 308 308 MANGANESE 2 (BY SIMILARITY).
FT METAL 347 347 MANGANESE 2 (BY SIMILARITY).
SQ SEQUENCE 383 AA; 43569 MW; 88E00C2BDAE4B9BF CRC64;

Query Match 5.2%; Score 134.5; DB 1; Length 383;
Best Local Similarity 22.4%; Pred. No. 0.072;
Matches 59; Conservative 41; Mismatches 92; Indels 71; Gaps 10;

OY 63 FLIVGFNGYDGNRVNPFVAQRLSALLGQLNAEHAEDVRVLLQAFDV----- 112
DB 86 FFYGLFDHGGTECESEFLSTNIGKIIENQDLN-----DTEKILKEVHSVGGYNAGLKP 139
OY 113 -VERGFLESIDDALEAKASLSQSLPEGVFOHQLPPQYOKILERLKTLEK-----EISGGAM 167
DB 140 FSLRTVLQGRBDDLMRARLYSF-----LQADMDYLNTVAPSPDSAVPGAV 187
OY 168 AVVAVLNNK-----LVANVGTRALLCKSTVDGLQVTLNVDHTTENEDLEF 216
DB 188 GTVAITTSKNNSYWESDSYIHLAVGDTALLCDSRTG--RAHRLTFQHHPADVEEAR 245
OY 217 RLSQGLDAGKIKQVGIICGF-----STRIGD-YKVKGYTIDILLSAKSKPIIAE 269
DB 246 RLRRYVMGFSRDS-----FGQRFAMVAVNTRSFSGYKL-----KGLGVYAE 287
OY 270 PEIHGAQPLDGVTFVLMSGL 292
DB 288 PQLTSIHSILRDMSTFLTLSDGI 310

```

Search completed: December 9, 2002, 22:48:23
 Job time : 23.8726 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 9, 2002, 18:36:04 ; Search time 25.1274 Seconds
(without alignments)
955.720 Million cell updates/sec

Title: US-09-830-144-2
Perfect score: 3014
Sequence: 1 MSTASAASSSSSSSAGEMIE.....QCKKQLEVRQQKQKQTS 579

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2986	99.1	579	1 M3K7 MOUSE	Q62073 mus musculu
2	2982.5	99.0	606	1 M3K7 HUMAN	Q43318 homo sapien
3	501.5	16.6	393	1 M3K7 DROME	P83104 drosophila
4	493.5	16.4	954	1 M3KA HUMAN	Q02779 homo sapien
5	418.5	13.9	394	1 M3K9 HUMAN	P80192 homo sapien
6	414	13.7	410	1 KYK2 DICDI	P18161 dictyosteli
7	412	13.7	859	1 M3KC HUMAN	Q12852 homo sapien
8	408.5	13.6	888	1 M3KC MOUSE	Q60700 mus musculu
9	408	13.5	888	1 M3KC RAT	Q63796 rattus norv
10	398	13.2	821	1 CTR1 ARATH	Q05609 arabidopsis
11	379.5	12.6	1584	1 KYK1 DICDI	P18160 dictyosteli
12	373	12.4	832	1 ANR3 HUMAN	P57078 homo sapien
13	365.5	12.1	630	1 TEC MOUSE	P24604 mus musculu
14	365.5	12.1	631	1 TEC HUMAN	P42680 homo sapien
15	363	12.0	2347	1 KR03 HUMAN	P08922 homo sapien
16	362.5	12.0	1068	1 FAK1 XENLA	Q91738 xenopus lae
17	361	12.0	1055	1 FAK1 RAT	Q53346 rattus norv
18	358.5	11.9	620	1 ITK HUMAN	Q08881 homo sapien
19	358.5	11.9	984	1 EPA3 RAT	Q08680 rattus norv
20	356.5	11.8	1052	1 FAK1 MOUSE	Q34152 mus musculu
21	356	11.8	625	1 ITK MOUSE	Q03526 mus musculu
22	354.5	11.8	527	1 TXK MOUSE	P24682 mus musculu
23	354.5	11.8	983	1 EPA3 HUMAN	P29320 homo sapien
24	353.5	11.7	1114	1 RET HUMAN	P07949 homo sapien
25	352	11.7	1115	1 RET MOUSE	P35346 mus musculu
26	349.5	11.6	1052	1 FAK1 HUMAN	Q05397 homo sapien
27	349	11.6	380	1 KMIL AVIMH	P00531 avian retro
28	347.5	11.5	981	1 EPA3 BRARE	O13146 brachydanio
29	346.5	11.5	804	1 FPS DROME	P18106 drosophila
30	345.5	11.5	1620	1 ALK HUMAN	Q9um73 homo sapien
31	344.5	11.4	983	1 EPA3 CHICK	P29318 gallus gall
32	344.5	11.4	1037	1 EPA5 HUMAN	P54756 homo sapien
33	343	11.4	467	1 MATK RAT	P41243 rattus norv

RESULT 1					
M3K7 MOUSE					
ID M3K7 MOUSE	STANDARD;	PRT;	579 AA.		
AC Q62073;					
DT 16-OCT-2001 (Rel. 40, Created)					
DT 16-OCT-2001 (Rel. 40, Last sequence update)					
DT 16-OCT-2001 (Rel. 40, Last annotation update)					
DE Mitogen-activated protein kinase kinase 7 (EC 2.7.1.-)					
DE (Transforming growth factor-beta-activated kinase 1) (TGF-beta-activated kinase 1).					
DE MAP3K7 OR TAK1.					
GN MAP3K7 OR TAK1.					
OS Mus musculus (Mouse).					
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.					
OC NCBI_TaxID=10090;					
RN [1]					
RP SEQUENCE FROM N.A.					
RX MEDLINE=96123277; PubMed=8533096;					
RA Yamaguchi K., Shirakabe K., Shibuya H., Irie K., Ohishi I., Ueno N., Taniguchi T., Nishida E., Matsumoto K.;					
RT "Identification of a member of the MAPKKK family as a potential mediator of TGF-beta signal transduction.";					
RL Science 270:2008-2011(1995).					
CC -!- FUNCTION: CAN PHOSPHORYLATE AND ACTIVATE YET UNDEFINED MAPKKS.					
CC -!- MEDIATOR OF TGF-BETA SIGNAL TRANSDUCTION. STIMULATES NF-KAPPA B ACTIVATION.					
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.					
CC MAP KINASE KINASE SUBFAMILY.					
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; STRONG, TO C21ORF7.					
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).					
CC EMBL; D76446; BAA11184.1; -.					
CC HSSP; P08631; 1AD5.					
CC MGD; MGI:1346877; Map3k7.					
CC InterPro; IPR000719; Euk_pkinase.					
CC InterPro; IPR004040; STY_pkinase.					
CC InterPro; IPR002290; Ser_thr_pkinase.					
CC Pfam; PF00069; pkinase; 1.					
CC ProDom; PD000001; Euk_pkinase; 1.					
CC SMART; SM00221; STYKc; 1.					
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.					
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.					
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.					
CC Transferrase; Serine/threonine-protein kinase; ATP-binding.					
FT DOMAIN 8 16 POLY-SER.					
FT NP_BIND 36 291					
FT BINDING 42 50					
FT ACT_SITE 63 63					
FT BY SIMILARITY 156					

P05625 gallus gall
P29319 mus musculu
P08630 drosophila
O61772 mus musculu
O63572 rattus norv
P16092 mus musculu
Q62956 rattus norv
P54755 gallus gall
Q922p5 rattus norv
P11362 homo sapien
P54761 mus musculu
P58500 mus musculu

ALIGNMENTS

```

SQ SEQUENCE 579 AA; 64227 MW; 97C9F6F3C8E283EE CRC64;
Query Match 99.1%; Score 2986; DB 1; Length 579;
Best Local Similarity 99.1%; Pred. No. 9,6e-168;
Matches 574; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSTASAAASSSSSSAGEMIEAPSOVLNFEIDYKEIEVEEVGRGAFGVCAKAKRAKOV 60
DB 1 MSTASAAASSSSSSAGEMIEAPSOVLNFEIDYKEIEVEEVGRGAFGVCAKAKRAKOV 60
QY AIKOIESESEKKAFFIVELROLSRVNHNPVLYGACLNIPVCLMEYAEGSLYVNLHGAE 120
DB AIKOIESESEKKAFFIVELROLSRVNHNPVLYGACLNIPVCLMEYAEGSLYVNLHGAE 120
QY PLPYTYAAHAMSWCLOCSQGVAYLHSMOPKALIHRLKPPNLLVAGGVTLKICDGTAC 180
DB PLPYTYAAHAMSWCLOCSQGVAYLHSMOPKALIHRLKPPNLLVAGGVTLKICDGTAC 180
QY DIQHTMTNNKGSAAAMAPVEFEGSNSEKCDVFSWGIIMEVITRRKPPDEIGGPAFRIM 240
DB DIQHTMTNNKGSAAAMAPVEFEGSNSEKCDVFSWGIIMEVITRRKPPDEIGGPAFRIM 240
QY 181 DIQHTMTNNKGSAAAMAPVEFEGSNSEKCDVFSWGIIMEVITRRKPPDEIGGPAFRIM 240
DB 181 DIQHTMTNNKGSAAAMAPVEFEGSNSEKCDVFSWGIIMEVITRRKPPDEIGGPAFRIM 240
QY 241 WAVHNGTRPPLINKLPKPIESLMTRCMSKDPQSPRPMEEIVKIMTHLMRYFGADEPLQY 300
DB 241 WAVHNGTRPPLINKLPKPIESLMTRCMSKDPQSPRPMEEIVKIMTHLMRYFGADEPLQY 300
QY 301 PCOYSDGGSNSATSTGSEFMDIASTNTSNKSDTNMEQVATNDTIKRLSKLLKNOAKQ 360
DB 301 PCOYSDGGSNSATSTGSEFMDIASTNTSNKSDTNMEQVATNDTIKRLSKLLKNOAKQ 360
QY 361 SESGRSLGASHSSVESLPTSEGRKMSADMSIEFARIAATTGNOQPRRSIOQLTVYG 420
DB 361 SESGRSLGASHSSVESLPTSEGRKMSADMSIEFARIAATTGNOQPRRSIOQLTVYG 420
QY 421 TEBGQVSSSSSPSVAMITTSQPTSEKPTRSHFWTPDSDTDNGSDNSIPMAYLTLDHQL 480
DB 421 TEBGQVSSSSSPSVAMITTSQPTSEKPTRSHFWTPDSDTDNGSDNSIPMAYLTLDHQL 480
QY 481 QPLAPCPNKSKEHNAVEFHCKMAQOBYMKVQTEITALLQKQELVALLDQEDQDQNTSRL 540
DB 481 QPLAPCPNKSKEHNAVEFHCKMAQOBYMKVQTEITALLQKQELVALLDQEDQDQNTSRL 540
QY 541 VOEHKKLIDENKSLSTYYOOCKKOLEVIRSQOQKRGTS 579
DB 541 VOEHKKLIDENKSLSTYYOOCKKOLEVIRSQOQKRGTS 579

RESULT 2
M3K7 HUMAN
ID _M3K7 HUMAN STANDARD; PRT; 606 AA.
AC O43318; O43317; O43319;
DT 16-OCT-2001 (Rel. 40; Created)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 15-JUN-2002 (Rel. 41; Last annotation update)
DE Mitogen-activated protein kinase kinase 7 (BC 2.7.1.-)
DE (Transforming growth factor-beta-activated kinase 1) (TGF-beta-
DE activated kinase 1).
GN Homo sapiens (Human).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euteria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxId=9606;
RN NCBI_TaxId=9606;
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC TISSUE=Lung;
RC MEDLINE=98153801; PubMed=9480845;
RA Sakurai H., Shigemori N., Hasegawa K., Sugita T.;
RT "TGF-beta-activated kinase 1 stimulates NF-kappa B activation by an
RT NF-kappa B-inducing kinase-independent mechanism.";
RL Biochem. Biophys. Res. Commun. 243:545-549 (1998).
CC -I- FUNCTION: CAN PHOSPHORYLATE AND ACTIVATE YET UNDEFINED MAPKKS.
CC MEDIATOR OF TGF-BETA SIGNAL TRANSDUCTION. STIMULATES NF-KAPPA B
CC ACTIVATION.
```

```

CC -I- ALTERNATIVE PRODUCTS: 3 ISOFORMS, 1A, 1B (SHOWN HERE) AND 1C; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -I- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE SUBFAMILY.
CC -I- SIMILARITY: IN THE C-TERMINAL SECTION, STRONG, TO C21ORF7.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB009357; BAA5026.1; -
DR EMBL; AB009356; BAA5025.1; -
DR EMBL; AB009358; BAA5027.2; -
DR HSSP; P08631; 1AD5
DR Genew; HGNC; 6859; MAP3K7.
DR MTW; 602614; -
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR004040; STY_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase.1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00221; STYKc; 1.
DR PROSITE; PS00107; PROTEIN KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN KINASE_ST; 1.
DR PROSITE; PSS0011; PROTEIN KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Alternative splicing.
FT DOMAIN 8 14 POLY-SER.
FT NP_BIND 36 291 PROTEIN KINASE.
FT BINDING 42 50 ATP (BY SIMILARITY).
FT ACT_SITE 63 63 ATP (BY SIMILARITY).
FT ACT_SITE 156 156 BY SIMILARITY.
FT VARSPIC 404 430 MISSING (IN ISOFORM 1A).
FT VARSPIC 509 518 PLAPCNSKE -> ARTSCTRTPG (IN ISOFORM 1C).
FT VARSPIC 519 606 MISSING (IN ISOFORM 1C).
SQ SEQUENCE 606 AA; 67196 MW; 3DBF8147CD174013 CRC64;

Query Match 99.0%; Score 2982.5; DB 1; Length 606;
Best Local Similarity 95.4%; Pred. No. 1.6e-167;
Matches 578; Conservative 0; Mismatches 1; Indels 27; Gaps 1;

QY 1 MSTASAAASSSSSSAGEMIEAPSOVLNFEIDYKEIEVEEVGRGAFGVCAKAKRAKOV 60
DB 1 MSTASAAASSSSSSAGEMIEAPSOVLNFEIDYKEIEVEEVGRGAFGVCAKAKRAKOV 60
QY 61 AIKOIESESEKKAFFIVELROLSRVNHNPVLYGACLNIPVCLMEYAEGSLYVNLHGAE 120
DB 61 AIKOIESESEKKAFFIVELROLSRVNHNPVLYGACLNIPVCLMEYAEGSLYVNLHGAE 120
QY 121 PLPYTYAAHAMSWCLOCSQGVAYLHSMOPKALIHRLKPPNLLVAGGVTLKICDGTAC 180
DB 121 PLPYTYAAHAMSWCLOCSQGVAYLHSMOPKALIHRLKPPNLLVAGGVTLKICDGTAC 180
QY 181 DIQHTMTNNKGSAAAMAPVEFEGSNSEKCDVFSWGIIMEVITRRKPPDEIGGPAFRIM 240
DB 181 DIQHTMTNNKGSAAAMAPVEFEGSNSEKCDVFSWGIIMEVITRRKPPDEIGGPAFRIM 240
QY 241 WAVHNGTRPPLINKLPKPIESLMTRCMSKDPQSPRPMEEIVKIMTHLMRYFGADEPLQY 300
DB 241 WAVHNGTRPPLINKLPKPIESLMTRCMSKDPQSPRPMEEIVKIMTHLMRYFGADEPLQY 300
QY 301 PCOYSDGGSNSATSTGSEFMDIASTNTSNKSDTNMEQVATNDTIKRLSKLLKNOAKQ 360
DB 301 PCOYSDGGSNSATSTGSEFMDIASTNTSNKSDTNMEQVATNDTIKRLSKLLKNOAKQ 360
QY 361 SESGRSLGASHSSVESLPTSEGRKMSADMSIEFARIAATT----- 403
DB 361 SESGRSLGASHSSVESLPTSEGRKMSADMSIEFARIAATTATTAISKPRGRKTAISFN 420
```

```

Qy 404 -----GNGOPRRRSTODLTVTGTEPGQVSSSSSSPSVRMITTSPTSEKPTRSH 453
Db 421 ILDPVEIVISGNGOPRRRSTQDLTVTTEPGQVSSSSSSPSVRMITTSPTSEKPTRSH 480
Qy 454 WTPDDSTDTGNSNSIPMAYLTLDHQLOPLAPCNSKESMAVFEQCKWAEYMKVOTEI 513
Db 481 WTPDDSTDTGNSNSIPMAYLTLDHQLOPLAPCNSKESMAVFEQCKWAEYMKVOTEI 540
Qy 514 ALLLQKQELVAELDDQEKDQONTSLRVLQVHKLLDENKSLSYQOCKQLEIVRSQQ 573
Db 541 ALLLQKQELVAELDDQEKDQONTSLRVLQVHKLLDENKSLSYQOCKQLEIVRSQQ 600
Qy 574 KRGQTS 579
Db 601 KRGQTS 606

RESULT 3
M3K7_DROME STANDARD; PRT; 393 AA.
AC P83104;
AT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Putative mitogen-activated protein kinase kinase 7
DE (EC 2.7.1.-).
GN TAKL1.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
CX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazie R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.B.G.,
RA Abrial J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Fesler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mourtulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.J., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."

```

```

RL Science 287:2185-2195 (2000).
RN [2]
RP CONCEPTUAL TRANSLATION.
RA Manning G., Sudarsanam S., Plowman G.;
RT "Prediction of novel protein kinases from the Drosophila genome
RT project and EST sequences.";
RL Unpublished observations (AUG-2001).
CC -!- FUNCTION: Can phosphorylate and activate yet undefined MAPKKS.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE SUBFAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE003732; -; NOT ANNOTATED_CDS.
CC FlyBase; FBgn0046689; Takl1.
CC InterPro; IPR000719; Euk_Pkinase.
CC InterPro; IPR002290; Ser_Chtr_Pkinase.
CC InterPro; IPR001245; Tyr_Pkinase.
CC Pfam; PF00069; pkinase; 1.
CC PRINTS; PR00109; TYRKINASE.
CC ProDom; PD000001; Euk_pkinase; 1.
CC SMART; SM00220; S_TKC; 1.
CC SMART; SM00219; TyrKC; 1.
CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.
CC PROSITE; PS00111; PROTEIN KINASE DOM; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC KW Hypothetical protein; Transferase; Serine/threonine-protein kinase;
KW Tyrosine-protein kinase; ATP-binding.
FT DOMAIN 11 266 PROTEIN KINASE.
FT NP_BIND 17 25 ATP (BY SIMILARITY).
FT BINDING 38 38 ATP (BY SIMILARITY).
FT ACT_SITE 133 133 BY SIMILARITY.
SQ SEQUENCE 393 AA; 45238 MW; 43370C98490F0384 CRC64;
Query Match 16.6%; Score 501.5; DB 1; Length 393;
Best Local Similarity 31.1%; Pred. No. 1.3e-22;
Matches 133; Conservative 86; Mismatches 136; Indels 73; Gaps 16;
Qy 29 EIIDYKEIEV-EVVVGRGAFGVCKAKWRADVAIK-----QIESSESRKAFIVELRQ 80
Db 3 KQVDFAPNFKLSEKFGAGGAVRATFQNOEIAVKIPDFLEETIKKNAER-----EITH 57
Qy 81 LSRVNHPIVKLYGACLN--PVLVMEYAEGLSYLVNLHGAELPYPYTAHAHMSWCLQCS 138
Db 58 LSEIDHENVIRVIGRASNGKDYLLMEVLESGSLHNYLYGDDKWE-YTVEQAVRWALQCA 116
Qy 139 QGVAYLHNSQKALIHRLDKPNNLLLVAGGTVLKICDPTACDICTHMTNNKSGSAAMWAP 198
Db 117 KALAYLHSLD-RPIVHRDIKQNMLLYNQHEDLKICDPLATDMNSNKTDMQGTLYMAP 175
Qy 199 EVFEGSNSEKCDVPSGCIILWEVITRKPDEITGGP--AFRIMWVHNGTRPPL---IK 253
Db 176 EAIKHLKYTAKDVSFYGIMLWEIMLTQPLSHLENPNISQYAIMKAISSEKGLPMEAVRS 235
Qy 254 NLPKPIESLIMTRCNSKDPSPSMBEIVKIMTHLMRYPPGADEPLQPCQYSDGQSNSA 313
Db 236 DCFEGIKQLMECCMDINPEKSPSKMEIKFGE--QYESGTDEDFIKPL---DEDTVAV 290
Qy 314 T----STGSFMDIASTNTSNKSDTNEQVPATNDTI-----RKLESKLLKNAQKQOSE 362
Db 291 TVHVDSSGSR-----MRVDFWRHQLPSIRMTFPIVKREARLGTVVREMAKAAAD 342
Qy 363 SGRSLGASHGSSVESLPPTSEGKMSADMSEIEARIAATGTNGQPPRRRSTQDLTVTGT 422
Db 343 GDR-----EVRRAEKD-TERETSRAHNGERETRR-----AGQD 375
Qy 423 PQQVSSRS 430

```



```
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Mitogen-activated protein kinase kinase 9 (BC 2.7.1.1.-) (Mixed
DE lineage kinase 1) (Fragment).
GN MAP3K9 OR MLK1 OR PRKEL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon epithelium;
RX MEDLINE=93238756; PubMed=8477742;
RA Dorow D.S., Devereux L., Dietzsch E., de Kretser T.;
RT "Identification of a new family of human epithelial protein kinases
RT containing two leucine/isoleucine-zipper domains.";
RL Eur. J. Biochem. 213:701-710(1993).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN EPITHELIAL TUMOR CELL LINES OF
CC COLONIC BREAST AND OESOPHAGEAL ORIGIN.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -1- MAP KINASE KINASE SUBFAMILY.
DR PIR; S32467; S32467.
DR PIR; JU0229; JU0229.
DR HSSP; P12931; LFMK.
DR GENE; HGNC:6861; MAP3K9.
DR MIM; 600136; -.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
KW Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase;
KW ATP-binding.
FT NON_TER 1 1
FT DOMAIN 3 271 PROTEIN KINASE.
FT NP_BIND 9 17 ATP (BY SIMILARITY).
FT BINDING 30 30 ATP (BY SIMILARITY).
FT ACT_SITE 127 127 BY SIMILARITY.
FT DOMAIN 289 310 LEUCINE-ZIPPER 1 (BY SIMILARITY).
FT DOMAIN 324 345 LEUCINE-ZIPPER 2 (BY SIMILARITY).
FT DOMAIN 354 368 ARG/LYS-RICH (BASIC).
FT SEQUENCE 394 AA; 44975 MW; DBE40B7D31047FD8 CRC64;

Query Match 13.9%; Score 418.5; DB 1; Length 394;
Best Local Similarity 36.8%; Pred. No. 9.3e-18;
Matches 98; Conservative 47; Mismatches 100; Indels 23; Gaps 8;

QY 35 EIEVEVVGAFGVCKAKRAKDVAIK--QTESERKAFIVELRQ----LSRVNHPN 88
DB 2 ELTLEEIIIGGKGVTRAFWGDVAVKARHDPDEDISQTIENVQKELFAMLKHPN 61
QY 89 IVKLYGACLPN--VCLVMEYAEGLSLNVLHGAPLPFYTTAAHMSWCLQCSQVAYLHS 146
DB 62 IIALRGVCLKEPNLCLVMEFARGGSLNVLGSKRIPDI---LVNVAVQIARGMYLHD 117
QY 147 MQPKALIHRLDKPNNLLV-----AGTVLKICDFTGACD-IQTHMTNKGSAAMAP 198
DB 118 EAIVPPIIHRDLKSSNIILOKVENGDLSNLIKLTIDTGLAREWHRTTQNSAAGTYAMAP 177
QY 199 EVFEGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMAV-HNGTRPPLIKNLPK 257
DB 178 EVTRASFSGKSDVMSVGVLLWELLTGEVFPFGIDG--LRVAVGVAMNKALPIPTCBE 235
QY 258 PIESLMTRCWSKDPSPQSPSEIEIVKIMT 285
DB 236 PFAKLMDOWNPDHPHSRPSFTNILDQLT 263

RESULT 6
QYK2_DICDI
```

QY 304 YSDEGQSNATSTGSEFMDISTNTSKSDT 333
 Db 376 VSDQFYQYSEPTPR--LALSNQSSNSS 402

RESULT 7
 M3KC_HUMAN STANDARD; PRT; 859 AA.

AC 012852;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Mitogen-activated protein kinase kinase kinase 12 (EC 2.7.1.37)
 DE (leucine-zipper protein kinase) (ZPK).
 GN MAP3K12 OR ZPK.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RN 11
 RP SEQUENCE FROM N.A.
 RC TISSUE=Teratocarcinoma;
 RX MEDLINE=94311945; PubMed=8037767;
 RA Reddy U.R., Pleasure D.,
 RT "Cloning of a novel putative protein kinase having a leucine zipper
 domain from human brain.";
 RL Biochem. Biophys. Res. Commun. 202:613-620(1994).
 CC -1- FUNCTION: May be an activator of the JNK/SAPK pathway.
 CC Phosphorylates beta-casein, histone 1 and myelin basic protein in
 CC vitro.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -1- COFACTOR: Magnesium.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated (By
 CC similarity).
 CC -1- TISSUE SPECIFICITY: Highly expressed in brain and kidney.
 CC -1- PTM: Autophosphorylated on Ser/Thr. Phosphorylated in cytosol
 CC under basal conditions and dephosphorylated when membrane-
 CC associated (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC MAP KINASE KINASE KINASE SUBFAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).

 CC EMBL: U07358; AAA67343.1; -
 DR HSSP: P12931; 1FMK.
 DR GeneW: HGNC:6851; MAP3K12.
 DR MIM: 600447; -
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR004040; STY_pkinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR Pfam: PF00069; pkinase; 1.
 DR ProDom: PD000001; Euk_pkinase; 1.
 DR SMART: SM00221; STYKC1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP, FALSE_NEG.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
 KM Transferrase; Serine/threonine-protein kinase; ATP-binding;
 KM Phosphorylation; Magnesium; Membrane.
 FT DOMAIN 125 366 PROTEIN KINASE.
 FT NP_BIND 131 139 ATP (BY SIMILARITY).
 FT BINDING 152 152 ATP (BY SIMILARITY).
 FT ACT_SITE 236 236 BY SIMILARITY.
 FT DOMAIN 665 668 POLY-PRO.
 FT DOMAIN 720 725 POLY-GLU.
 SQ SEQUENCE 859 AA; 93188 MW; 0E5209792C5C6F05 CRC64;

Query Match 13.7%; Score 412; DB 1; Length 859;
 Best Local Similarity 24.6%; Pred. No. 5,9e-17;
 Matches 133; Conservative 79; Mismatches 181; Indels 148; Gaps 20;

QY 30 EIDYKEIEVEEVRGAGVGVCKAKRAKDAVAIKQISESESRKAFIVELRQLSVNPN 89
 Db 119 EVFPEELDLQWVGSGAGVFLGRFHGEYAVAKVDLKE-----IDIKLKKLKPN 173
 QY 90 VKLYGACLP--VCLVMEYAGSLVYVLHGAEDLPYTAHAAMSWCLQSGVAYLHSM 147
 Db 174 ITFGKGVCTQAPCCYLMEFCAQGLYEVLRAGRNV--TSSLVDVDMGAGGNVYLHL 230
 QY 148 QPKALIHDLKPPVLLVAGTVLKIDFGACDIOQHMTNNK--GSAAMABRVFEGSN 205
 Db 231 K--IHRDLKSPN-MLITVDVVKISDFGTSKSLSDKSTKMSFAGTVAWMAPEVIRNEP 286
 QY 206 YSEKDVFSWGIIMVEVITRRKPFDEIGGPAFRIMAV-HNGTPPPLIKLPKPIESLMT 264
 Db 287 VSEKVDIMSFQVLMELLTGELPKVDSSA--IIVGSGNSLHLVPSSCPDGFILLR 344
 QY 265 RCWSKDPSPQRPSEMEIVKIMTHLMRPYPGADPLQPCQYSDGQSNATSTGSFMDIAS 324
 Db 345 QCMNSKPRNRPSPFQ--ILLHL-----DIAS 368
 QY 325 TNT-SNKSdT-----NMQVPATNDITRLSKLL----- 353
 Db 369 ADVLSTPQETVFKSQAEWREVKLHFEEKISEGTCLHREELVMRRRELRLALDIRH 428
 QY 354 -----KNOAKQSSSGRL-----SLGASGSSVES 378
 Db 429 YERKLERANNIMELNMLQLELKERELLRREVALERRCPGLIKPSPGLHGNMEX 488
 QY 379 L-----PTSEGRKMSADNSEIEARI-----AATNGQPRRSIODLTVTGPEGVSS 428
 Db 489 LIKKRNVPQNLSPHQRPDLKAEKSLPKLDALSLGVLP-----GCPKAPSPGR--S 540
 QY 429 RSSPSVAMITTSPTSEKP---TRSHPTPDDSTDING-----SUNSLPMAYLTIDHQ 479
 Db 541 RRGKTRHRKSAKSCDPLRLTAVPHPHGGPSPGGLGGPSAWACPPALRGLLHD 600
 QY 480 L 480
 Db 601 L 601

RESULT 8
 M3KC_MOUSE STANDARD; PRT; 888 AA.

AC Q60700; P70286;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Mitogen-activated protein kinase kinase kinase 12 (EC 2.7.1.37)
 DE (leucine-zipper protein kinase) (ZPK) (Dual leucine zipper bearing
 DE kinase) (DLK).
 GN MAP3K12 OR ZPK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=CD-1; TISSUE=Brain;
 RX MEDLINE=95074107; PubMed=7963011;
 RA Holzman L.B., Meritt S.B., Fan G.,
 RT "Identification, molecular cloning, and characterization of dual
 RT leucine zipper bearing kinase. A novel serine/threonine protein kinase
 RT that defines a second subfamily of mixed lineage kinases.";
 RL J. Biol. Chem. 269:30808-30817(1994).
 RN 12
 RP SEQUENCE FROM N.A.
 RC STRAIN=ICR X Swiss Webster; TISSUE=Brain;
 RX MEDLINE=96365386; PubMed=8769565;

RA Blouin R., Beaudoin J., Bergeron P., Nadeau A., Grondin G.;
 RT "Cell-specific expression of the ZPK gene in adult mouse tissues.";
 RL DNA Cell Biol. 15:631-642(1996).
 RN [3]
 RP PHOSPHORYLATION, AND MUTAGENESIS OF LYS-185 AND GLU-192.
 RX MEDLINE=96279269; PubMed=8663324; Mol Cell Biol. 15:192-197.
 RA Mata M., Merritt S.E., Fan G., Yu G.G., Holzman L.B.;
 RT "Characterization of dual leucine zipper-bearing kinase, a mixed
 RT kinase present in synaptic terminals whose phosphorylation
 RT state is regulated by membrane depolarization via calcineurin.";
 J. Biol. Chem. 271:16888-16896(1996).
 CC -!- FUNCTION: May be an activator of the JNK/SAPK pathway.
 CC Phosphorylates beta-casein, histone 1 and myelin basic protein in
 CC vitro.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -!- COFACTOR: Magnesium.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated.
 CC -!- TISSUE SPECIFICITY: Expressed in brain, kidney, lung, heart,
 CC testis, gastrointestinal tract, stomach, liver and pancreas.
 CC Within the nervous system, predominantly expressed in neurons and
 CC enriched in synaptic terminals.
 CC -!- PTM: Autophosphorylated on Ser/Thr. Phosphorylated in cytosol
 CC under basal conditions and dephosphorylated when membrane-
 CC associated.
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC MAP KINASE KINASE SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U14636; AA57280.1; --
 CC EMBL; U23789; AAB17123.1; --
 CC HSPSP; P12931; IFMK.
 DR MGD; MGI:1346881; Mapk12.
 DR InterPro; IPR000719; Euk_pkinase.
 DR InterPro; IPR004040; STY_pkinase.
 DR Pfam; PF00669; pkinase; 1.
 DR ProDom; PD000001; Euk_pkinase; 1.
 DR SMART; SM00221; STYK; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 KW Transferase, Serine/threonine-protein kinase; ATP-binding;
 KW Phosphorylation; Magnesium; Membrane.
 FT DOMAIN 158 399 PROTEIN KINASE.
 FT NP_BIND 164 172 ATP (BY SIMILARITY).
 FT BINDING 185 185 ATP.
 FT ACT_SITE 269 269 BY SIMILARITY.
 FT DOMAIN 56 62 POLY-GLY.
 FT DOMAIN 668 671 POLY-PRO.
 FT DOMAIN 698 701 POLY-GLU.
 FT DOMAIN 753 758 POLY-GLU.
 FT MUTAGEN 185 185 K->A: NO CATALYTIC ACTIVITY.
 FT MUTAGEN 192 192 E->A: NO CHANGE.
 FT CONFLICT 18 18 V->A (IN REF. 2).
 FT CONFLICT 28 29 KL->NV (IN REF. 2).
 FT CONFLICT 382 382 S->T (IN REF. 2).
 FT CONFLICT 494 495 EQ->DE (IN REF. 2).
 FT CONFLICT 517 517 N->D (IN REF. 2).
 FT CONFLICT 794 794 E->G (IN REF. 2).
 SQ SEQUENCE 888 AA; CFECFD134F889ABB CRC64;
 Query Match 13.6%; Score 408.5; DB 1; Length 888;
 Best Local Similarity 24.9%; Pred. No. 9.9e-17;
 Matches 125; Conservative 76; Mismatches 157; Indels 145; Gaps 18;
 QY 30 EIDYKIEVEVVGAGFVGVCKAKWRAKDVAIKQIESESEKAFIVELRQLSRVNHPI 89

Db 152 EYPFEILLQWVGSGAGAVFLGRFHGEVAVKVRDLKE-----TDIKHLRKLKHPNI 206
 QY 90 VKLYGACINP--VCLVMEYAGGSLYNVLHGAELPLPYTAAHAMSWCLQCQGVAYLHSM 147
 Db 207 ITFGKGVCTQAPCYCILMEFCAQQLYELRAGRPV---TPSLLDVMSMGIAAGMNYLHL 263
 QY 148 QPKALIHRLDKPPNLLVAGGTVLKICDFGTACDIQTHMTNKK--GSAAMWAPVFECSN 205
 Db 264 K---IIHRDLKSPN-MLITYDDVVKISDFGTSKELSDKSTKMSFAGTVAWVAPEVIRNEP 319
 QY 206 YSEKCDVFSWGIILMEVITRRKPFDEIGGPAFRIMWAV-HNGTRPPLIKNLPKPIESLMT 264
 Db 320 VSEKVDIWSFGVWLWELLTGEIPKVDVSSA--IIVGVSNSLHLPVSSCPDGFKILLR 377
 QY 265 RWSKDKPQRORSMEIEVKIMTHLMRYFFGCADEPLQPCQYDDEGQSNSATSGSFMDIAS 324
 Db 378 QCNWSKPNRPFSPQ---ILLHL-----DIAS 401
 QY 325 TNT-SNKSDT-----NMEQVATNDTIKRLESKLL----- 353
 Db 402 ADVLSTPQETVFKSOAWEREVVKLHFEKIKSEGTCILHLEELVWRRREELRHALDIREH 461
 QY 354 -----KNOAKQOSESGR-----SLGASHGSSVES 378
 Db 462 YERKLERANNLYMELNALMLQELKERELLRREQALERRCPGLKSHPSRCLLHGNTMEK 521
 QY 379 L-----PPTSEGRKMSADMSEIEARI-----AATGNGOPRRRSIQDLTVTGTSPGQVS- 427
 Db 522 LIKRNVFQKLSHPSKRPDLKTESLPLKDALSGVGLP-----GCPKGPSP 570
 QY 428 --SRSSSSPSVRMITTSGPTSEKP 448
 Db 571 GSRRGKTRHRKASAKSGCDLP 593
 RESULT 9
 M3KC RAT
 ID M3KC RAT STANDARD; PRT; 888 AA.
 AC Q63756;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Mitogen-activated protein kinase kinase kinase 12 (BC 2.7.1.37)
 DE (MAPK-upstream kinase) (MUK).
 GN MAP3K12 OR MUK.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96226099; PubMed=8637721;
 RA Hirai S., Izawa M., Osada S., Spyrou G., Ohno S.;
 RT "Activation of the JNK pathway by distantly related protein kinases,
 RT MEKK and MUK.";
 RL Oncogene 12:641-650(1996).
 CC -!- FUNCTION: May be an activator of the JNK/SAPK pathway.
 CC Phosphorylates beta-casein, histone 1 and myelin basic protein in
 CC vitro.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -!- COFACTOR: Magnesium.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated (By
 CC similarity).
 CC -!- PTM: Autophosphorylated on Ser/Thr. Phosphorylated in cytosol
 CC under basal conditions and dephosphorylated when membrane-
 CC associated (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC MAP KINASE KINASE SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D49785; BAA08621.1; -.
 DR HSP; P12911; IFMK.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR004040; STY_pkinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Euk_pkinase; 1.
 DR SMART; SM00221; STYK; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; FALSE_NEG.
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
 DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
 KM Transferase; Serine/threonine-protein kinase; ATP-binding;
 KW Phosphorylation; Magnesium; Membrane.
 FT DOMAIN 158 399 PROTEIN KINASE
 FT NP BIND 164 172 ATP (BY SIMILARITY).
 FT BINDING 185 185 ATP (BY SIMILARITY).
 FT ACT SITE 269 269 BY SIMILARITY.
 FT DOMAIN 56 62 POLY-GLY.
 FT DOMAIN 668 671 POLY-PRO.
 FT DOMAIN 698 701 POLY-PRO.
 FT DOMAIN 753 758 POLY-GLU.
 SQ SEQUENCE 888 AA; 96307 MW; 52AD964006BAE149 CRC64;
 Query Match 13.5%; Score 408; DB 1; Length 888;
 Best Local Similarity 28.3%; Pred.No. 1,1e-16;
 Matches 99; Conservative 67; Mismatches 114; Indels 70; Gaps 12;
 QY 30 EIDYKEIEVEEVGRFAFVCAKWRADVAIKQISESESEKAFIVEDRLSRVHPNI 89
 DB 152 EVFEELDLQWVGSGAQGVFLGRHGEVAVKVRDLKE-----TDIKHLRKXKHPNI 206
 QY 90 VKLYGACLPN--VCLVMEVAGSLYNVLHGAEPLPYRAAHAMSCVLCOSQGVAYLHSM 147
 DB 207 ITFKGVCTQAPCYCILMEFCAQQLYEVLRAGRPV---TPSLVDVSMGIAGMNVLLHIA 263
 QY 148 QPKALLHRDLKPNLLVAGCTYALKCDPCTACDICTHTMNNK--GSAAMMAPEVEGNS 205
 DB 264 K--IHRDKSPN-MLITYDDVAKISDFETSKELSDKSTKNSFAGTVAMMAPEVARNRP 319
 QY 206 YSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMAV-HNGTRPLINKLPPIESLMT 264
 DB 320 VSEKVDWMSGVVLWELLTGEIPIYKVDSSA--IIMGVGSNSLHLVPSSCPDGFILLR 377
 QY 265 RCWSKDPSPQPSMEIEVIMTHLMRYFPAGADEPLQYPCOYSDGQNSATSTGSPFDIAS 324
 DB 378 QCNMRKPRNRPFRQ--ILLHL-----DIAS 401
 QY 325 TNF-SNKSDF-----NMEQVPATNDTIKRLESKLLKNQAKO 359
 DB 402 ADVLSTPQETTFYFKSQAEWREEVKLFKEIKNSEGTCLHRLLEEELVMRRRE 451
 RESULT 10
 CTRL_ARATH
 ID CTRL_ARATH STANDARD; PRT; 821 AA.
 AC Q05609;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Serine/threonine-protein kinase CTRL (EC 2.7.1.37).
 GN CTRL OR AT5G03730 OR F17C15_150.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eurosidia; Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=cv, Columbia; TISSUE=Seedling;
 RX MEDLINE=3161417; PubMed=8431946;
 RA Kieber J.J., Rothenberg M., Roman G., Feldmann K.A., Ecker J.R.,
 RT "CTRL, a negative regulator of the ethylene response pathway in
 Arabidopsis, encodes a member of the raf family of protein kinases";
 RL Cell 72:427-441(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv, Columbia;
 RX MEDLINE=21016721; PubMed=11130714;
 RA Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E.,
 RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,
 RA Konata M., Matsumoto M., Matsuno A., Muraki A., Nakayama S.,
 RA Nakazaki N., Naito K., Okumura S., Shino S., Takeuchi C., Wada T.,
 RA Matubae A., Yamada M., Yasuda M., Sato S., de la Bastide M.,
 RA Huang E., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,
 RA Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J.,
 RA Stoneking T., Pepin K., Spieth J., Sekhon M., Armstrong J., Becker M.,
 RA Belter E., Cordum H., Cordes M., Courtney L., Courtney W., Dante M.,
 RA Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Latreille P.,
 RA Leonard S., Meyer R., Mulvaney E., Ozerky P., Riley A., Strommatt C.,
 RA Wagner-McPherson C., Wollam A., Yeakum M., Bell M., Dedhia N.,
 RA Parnell L., Shah R., Rodriguez M., Hoon See L., Vil D., Baker J.,
 RA Kirchhoff K., Toth K., King L., Bahret A., Miller B., Marra M.,
 RA Martienssen R., McCombie W.R., Wilson R.K., Murphy G., Bancroft I.,
 RA Volckaert G., Wambolt R., Duesterhoeft A., Stiekema W., Pohl T.,
 RA Entian K.-D., Terry N., Hartley N., Bent E., Johnson S.,
 RA Langham S.-A., McCullagh B., Robben J., Gymonprez B., Zimmermann W.,
 RA Ramsperger U., Wedler R., Balke K., Wedler E., Peters S.,
 RA van Staveren M., Dirke W., Moolman P., Klein Lankhorst R.,
 RA Weitzenecker T., Bothe G., Rose M., Hauf J., Benneiser S., Hempel S.,
 RA Feldbusch W., Lamberth S., Villarroel R., Giehl J., Archiles W.,
 RA Bents O., Lemcke K., Kolesov G., Mayer K.F.X., Rued S., Schoof H.,
 RA Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Franz P.F.;
 RT "Sequence and analysis of chromosome 5 of the plant Arabidopsis
 thaliana";
 RL Nature 408:823-826(2000).
 CC -1- FUNCTION: ACTS AS A NEGATIVE REGULATOR IN THE ETHYLENE RESPONSE
 CC PATHWAY.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN BOTH SEEDLINGS AND ADULT PLANTS.
 CC -1- MISCELLANEOUS: CTRL MUTANTS DISPLAY ETHYLENE-TREATED PHENOTYPES,
 CC RESULTING IN PLANTS WITH SMALL, UNEXPANDED LEAVES AND WHOSE SEED
 CC COTYLEDON GROWTH IS IMPAIRED.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC M1/RAF SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L08789; AAA32779.1; -.
 DR EMBL; L08790; AAA32780.1; -.
 DR EMBL; AL162506; CAB82938.1; -.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR004040; STY_pkinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Euk_pkinase; 1.
 DR SMART; SM00221; STYK; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
 DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
 KM Transferase; Serine/threonine-protein kinase; ATP-binding.
 KW Phosphorylation; Magnesium; Membrane.
 FT DOMAIN 65 69 POLY-GLY.
 FT DOMAIN 135 141 POLY-GLY.
 FT DOMAIN 551 557 PROTEIN KINASE.
 FT NP BIND 557 565 ATP (BY SIMILARITY).
 FT BINDING 578 578 ATP (BY SIMILARITY).

DB 1561 ----EMEDQVSSSFAS 1572

RESULT 12
ANR3_HUMAN STANDARD; PRT: 832 AA.
ID ANR3_HUMAN STANDBY; PRT: 832 AA.
AC P57078; Q96KH0.
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Serine/threonine-protein kinase ANKRD3 (EC 2.7.1.-) (Ankyrin repeat domain protein 3) (PKC-delta-interacting protein kinase).
GN ANKRD3 OR D1K.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Fetal kidney, and Fetal lung;
RA Shimizu N., Kudoh J., Shibuya K.;
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20289799; PubMed=10830953;
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T., Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E., Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A., Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D., Reichwald K., Rump A., Schillabel M., Schudy A., Zimmermann W., Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S., Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E., Minoshima S., Shimizu N., Nordliek G., Hornischer K., Brandt P., Schafte M., Schoen O., Desario A., Reichelt J., Kauer G., Blecker H., Ramser J., Beck A., Klages S., Hennig S., Rieselmann L., Dagand E., Wehnreyer S., Borzym K., Gardiner K., Nizetic D., Francis F., Lehnach H., Reinhardt R., Yaspo M.-L.;
RT "The DNA sequence of human chromosome 21.";
RL Nature 405:311-319(2000).
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are produced by alternative splicing.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -1- SIMILARITY: CONTAINS 10 ANK REPEATS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-stb.ch/announce/> or send an email to license@isb-stb.ch).
CC -----
DB EMBL: AB047783; BAB56136.1; -
DB EMBL: AP001743; BAA95526.1; -
DB HSSP: P25963; I1KN.
DR Genew: HGNC:496; ANKRD3.
DR MIM: 605706; -
DR InterPro: IPR002110; ANK.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR004040; STY_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00069; ank; 10.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PR00109; TYRKINASE.
DR PRODOM: PD000001; Euk_pkinase; 1.
DR SMART: SM00248; ANK; 10.
DR SMART: SM00221; STYK; 1.
DR PROSITE: PS00107; PROTEIN KINASE_ATP; 1.
DR PROSITE: PS00108; PROTEIN KINASE_ST; 1.
DR PROSITE: PS50011; PROTEIN KINASE_DOM; 1.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE: PS50088; ANK_REPEAT; 9.

KW Transferase; Serine/threonine-protein kinase; ATP-binding; Repeat;
KW ANK repeat; Alternative splicing.
FT DOMAIN 22 286
FT REPEAT 485 514 ANK 1.
FT REPEAT 518 547 ANK 2.
FT REPEAT 551 580 ANK 3.
FT REPEAT 584 613 ANK 4.
FT REPEAT 617 647 ANK 5.
FT REPEAT 651 680 ANK 6.
FT REPEAT 684 713 ANK 7.
FT REPEAT 717 746 ANK 8.
FT REPEAT 750 780 ANK 9.
FT REPEAT 782 811 ANK 10.
FT NP_BIND 28 36
FT BINDING 51 51
FT ACT SITE 143 143
FT VARSPIC 278 325
FT CONFLICT 714 714
SQ SEQUENCE 832 AA; 91610 MM; 5DBFFPD5F04F7ECB CRC64;
Query Match 12.4%; Score 373; DB 1; Length 832;
Best Local Similarity 27.9%; Pred. No. 1,1e-14;
Matches 168; Conservativity 87; Mismatches 213; Indels 134; Gaps 35;
40 EVNGRGAFGVYCKAK---WAKQVAIKQIES-----ESERAFIVELRQLSRVNHPNIVKL 92
26 EKVSQGGGQYQYKRVHWMKTV-LATCSPLHAVDERRELEAKKEMAKFRYLLPV 84
93 YGACINPVCGLMVEAYEGSLYNNVHGAEPLPY---YTAAMSMCLOCSQGVAVLHSMOP 149
85 YGICREPGVLMETMERGSLLEKL-ASEPLPWDRFRIIH-----ETAGNPNFLHMAP 137
150 KALIHRIKIPNLLVAGTVYKICDFGTA-CDIQTH-----MTNKGSAAMNAP-VEE 202
138 -PLHLDLKIPANILLDAHYHV-KISDFGLAKNGLSHSHDLSMDGLFGITAYLVPERIRE 195
203 GSN-YSEKCDVFSMGILLMEVITRKPF-DEIGSPARIMMAVHNGNR---PLIKLPLK 257
196 KSRLEFDTHDVYSPAIYVIGVLYKKPFADKKN--ILHIVKVKKGRRPLPVCRARPR 253
258 P---IESIMTRCWGSDPSQSPSM-----BEIVKIMTHLM---RY-FPG----- 293
254 ACSHLIRLMQCMQGDPRVRRPTFGQNGLNGELIRQVLAALLPVYGRARRSGEGRLESEV 313
294 ---ADEPLQYPCQYSDGQ-----SNSATSTGSFMDIASTNTSNKSDTNMEQVPATNDT 344
314 IIRVTCPLSPQBITSETEDICEKPDDEVKETADLDVKS-----PEPRSEVVPV---R 365
345 IKRLSKLLKQNAQQSSGSLSGA-----SHGSSVESTPTSGKMSADMSSE 394
366 LKRASAPTFDNDYLSLBSLLSQLDGVSQAVGPELSSRSSSESKLPSSGSGKRLSG-VSS 424
395 IEARIAATTNGQ-----PRRSIQDLTVTGTPEQGVSSRSSSPVSMITTSQPTSEKPT 449
425 VD---SAFSSRGSLSLSEPRRPSSTDICTDYOQKKLV-----ALVSGDTSLMK 472
450 RSHPTPDSDTDTNGSDNSIPMAVLTLDHQLQPIAPC-----PNSKESMAVFEQCKM 502
473 ILQPDVDLALDSGAS-----LHLHAYAGQGECAKMLLNANPNLSNRGSTPLH--M 525
503 AOEYKQVTEIALLQKQELVAELDD-----EKQOQNSRLVQEHKKLLDENKS 553
526 AVE-RRRVGVELLLARKISVNAK-DEDDQWTLHFAAQNDESSTRL-----LLEKNAS 577
554 LS 555
578 VN 579
RESULT 13
TEC MOUSE
ID TEC MOUSE
AC P24604; STANDARD; PRT: 630 AA.

DE Tyrosine-protein kinase Tec (EC 2.7.1.112).
 GN TEC OR PCKTK4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Blood;
 RA MEDLINE=95019807; PubMed=7934162;
 RA Sato K., Mano H., Ariyama T., Inazawa J., Yazaki Y., Hirai H.;
 RT "Molecular cloning and analysis of the human Tec protein-tyrosine
 kinase";
 RT Leukemia 8:1663-1672(1994).
 RL Leukemia 8:1663-1672(1994).
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- TISSUE SPECIFICITY: HEMATOPOIETIC CELL LINES INCLUDING MYELOID,
 CC B-, AND T-CELL LINEAGES.
 CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. BTK
 CC SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 IS THE INITIATOR.
 CC -1- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
 CC WWW="http://www.infobiogen.fr/services/chronocancer/Genes/TECID75.html".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D29767; BAA06171.1; -.
 DR HSSP: 006187; 1B55.
 DR Gene: HGNC:11719; TEC.
 DR MIM: 600583; -.
 DR InterPro: IPR001562; BTK.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR001849; PH.
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001452; SH3.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam: PF00017; SH2; 1.
 DR Pfam: PF00018; SH3; 1.
 DR Pfam: PF00069; pkinase; 1.
 DR Pfam: PF00169; PH; 1.
 DR Pfam: PF00779; BTK; 1.
 DR PRINTS: PR00452; SH3DOMAIN.
 DR PRINTS: PR00402; TECBTKDOMAIN.
 DR PRINTS: PR00109; TYRKINASE.
 DR ProDom: PD000001; Euk_pkinase; 1.
 DR ProDom: PD000066; SH3; 1.
 DR ProDom: PD000093; SH2; 1.
 DR ProDom: PD000107; BTK; 1.
 DR SMART: SM00107; BTK; 1.
 DR SMART: SM00233; PH; 1.
 DR SMART: SM00252; SH2; 1.
 DR SMART: SM00325; SH3; 1.
 DR SMART: SM00219; TyRK; 1.
 DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE: PS00109; PROTEIN KINASE TYR; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS50002; SH3; 1.
 DR PROSITE: PS50003; PH_DOMAIN; 1.
 DR Transferrase; Tyrosine-protein kinase; ATP-binding; SH2 domain;
 KM SH3 domain; Phosphorylation.
 KW DOMAIN 4 111 PH.
 FT DOMAIN 179 239 SH3.
 FT DOMAIN 247 345 SH2.

FT DOMAIN 370 623 PROTEIN KINASE.
 FT NP_BIND 376 384 ATP (BY SIMILARITY).
 FT BINDING 398 398 ATP (BY SIMILARITY).
 FT ACT_SITE 489 489 BY SIMILARITY.
 FT MOD_RES 519 519 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 SO SEQUENCE 631 AA; 73629 MW; A55DECAFP91A9022 CRC64;
 Query Match 12.1%; Score 365.5; DB 1; Length 631;
 Best Local Similarity 34.7%; Pred. No. 2.1e-14;
 Matches 94; Conservative 47; Mismatches 107; Indels 23; Gaps 11;
 Oy 30 EIDPKIEIEVEVRGAGPVCYKAKWRAX-DVAIKOI-ESESERKAPVILRLQSLRVNHP 87
 Db 364 EINSSELTPEKELSSGLFGVIRGLKMPQYKAIKIRBSGAMCEEDFIEAKVMKLTLP 423
 Oy 88 NIVKLYGACL-NPVCILMEVBEGLSYLVNLAHAEPLPYTTAAHMSWCLQSGVAYLH 145
 Db 424 KLVQLYGVCTQCKRIYIVTFEMERGLNLFRLQRO--GHFSRVLLSMQDVGEGMEYL- 480
 Oy 146 SMQKALIHRLDKPNNLLVAGTVLKIQDFGTA---CDIQHTMNNKGSAAWMAPEVF 201
 Db 481 --ERNSEIHRDLAARNCLVSEAG-VKVSDFGMARYFLDQYTSSGAKFPVKCEPEVF 537
 Oy 202 EGSVSEKCDVFSWGIILWEVITR-RKPEIGGPAFRIMWAVHNGTR--PLINKLPK 257
 Db 538 NYSFSSKSDVMSGVLMVEVFTGKMPFEKYN--YEVVTWTRGRLYQPLASNY-- 593
 Oy 258 PIESIMTRCWSKPSQPSRSMEEIVKIMTHLM 288
 Db 594 -VEVWLRWCQEKREGRSPEDLRTIDELV 623
 RESULT 15
 KROS HUMAN STANDARD; PRT; 2347 AA.
 ID KROS HUMAN STANDARD; PRT; 2347 AA.
 AC P08922; O15368;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Proto-oncogene tyrosine-protein kinase ROS precursor (EC 2.7.1.112)
 DE (c-ros-1).
 GN ROS1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90280463; PubMed=2352949;
 RA Birchmeier C., O'Neill K., Riggs M., Wigler M.;
 RT "Characterization of ROS1 cDNA from a human glioblastoma cell line";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:4799-4803(1990).
 RN [2]
 RP SEQUENCE OF 1790-2259 FROM N.A.
 RX MEDLINE=87064611; PubMed=3023956;
 RA Matsushime H., Wang L.-H., Shibuya M.;
 RT "Human c-ros-1 gene homologous to the v-ros sequence of UR2 sarcoma
 RT virus encodes for a transmembrane receptorlike molecule";
 RL Mol. Cell. Biol. 6:3000-3004(1986).
 RN [3]
 RP SEQUENCE OF 1854-2245 FROM N.A.
 RX MEDLINE=87064625; PubMed=3785223;
 RA Birchmeier C., Birnbaum D., Waitches G., Faano O., Wigler M.;
 RT "Characterization of an activated human ros gene";
 RL Mol. Cell. Biol. 6:3109-3116(1986).
 CC -1- FUNCTION: THIS IS A PROBABLY A CELL GROWTH OR DIFFERENTIATION
 CC FACTOR RECEPTOR WITH A TYROSINE-PROTEIN KINASE ACTIVITY.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. INSULIN
 CC RECEPTOR SUBFAMILY.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; M34353; AAA60278.1; -;
DR EMBL; M33599; AAA60277.1; -;
DR EMBL; M33368; AAA60277.1; JOINED.
DR EMBL; M33591; AAA60277.1; JOINED.
DR EMBL; M33592; AAA60277.1; JOINED.
DR EMBL; M33593; AAA60277.1; JOINED.
DR EMBL; M33594; AAA60277.1; JOINED.
DR EMBL; M33595; AAA60277.1; JOINED.
DR EMBL; M33596; AAA60277.1; JOINED.
DR EMBL; M33597; AAA60277.1; JOINED.
DR EMBL; M33598; AAA60277.1; JOINED.
DR EMBL; M33880; AAA36580.1; ALT_TERM.
DR PIR; A25223; TVHURS.
DR PIR; A24421; TVHURT.
DR HSP; P08631; IADS.
DR Genew; HGNC:10261; ROS1.
DR MIM; 165020; -.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR000033; Ldl_receptor_rep.
DR InterPro; IPR002011; RTK_naseil.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00041; fn3; 7.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00060; FN3; 5.
DR SMART; SM00135; LX; 2.
DR SMART; SM00219; TyrcK; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.
DR PROSITE; PS00239; RECEPTOR TYR_KIN_II; 1.
DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
DR Transferase; Tyrosine-protein kinase; Receptor; Transmembrane;
KW Glycoprotein; ATP-binding; Phosphorylation; Proto-oncogene;
KW Signal.
FT CHAIN 1 27 POTENTIAL.
FT CHAIN 28 2347 PROTO-ONCOGENE TYROSINE-PROTEIN KINASE
FT ROS.
FT DOMAIN 28 1859 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1860 1882 POTENTIAL.
FT DOMAIN 1883 2347 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 1945 2222 PROTEIN KINASE.
FT NP_BIND 1951 1959 ATP (BY SIMILARITY).
FT BINDING 1980 1980 ATP (BY SIMILARITY).
FT MOD_RES 2114 2114 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CARBOHYD 52 52 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 114 114 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 123 123 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 324 324 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 396 396 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 471 471 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 607 607 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 628 628 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 706 706 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 714 714 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 732 732 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 939 939 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 961 961 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1015 1015 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1087 1087 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1090 1090 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1095 1095 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1211 1211 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1272 1272 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 1330 1330 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1458 1458 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1461 1461 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1474 1474 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1499 1499 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1565 1565 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1669 1669 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1715 1715 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1738 1738 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1808 1808 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 2213 2213 N-> D (IN REF. 2 AND 3).
FT CONFLICT 2228 2228 QC -> KS (IN REF. 2 AND 3).
FT CONFLICT 2246 2246 EDGVDICLNSDDIM -> KFDSEFEFSFRCTVN (IN
FT REF. 2).
SQ SEQUENCE 2347 AA; 263956 MW; E14F3DFD410C1D2A CRC64;

Query Match 12.0%; Score 363; DB 1; Length 2347;
Best Local Similarity 27.5%; Pred. No. 1.5e-13;
Matches 126; Conservative 73; Mismatches 196; Indels 64; Gaps 20;

QY 4 ASAASSSSSAGEMIEAPSOVLNFEEDYKEIEVEEVVGRGAFGVVCKAKWRAKD--- 59
DB 1913 AAGVGLANACYAIHTLPTQBEIENLPAPPREKLTIRLLLGSGAFGEVYEG--TAVDILGV 1970

QY 60 -----VAIKQIE---SESERKAFIVELRQLSRVNHNPVVKLYGACL--NPVCLVMEYAE 108
DB 1971 GSGEIKVAVKTKKSGTDQEKIEFLKEAHLMSKFNHPNLIKQLGVCLLNEPOYIILELME 2030

QY 109 GGSLYNVLHGAEPYPIY-----TAAHAMSWCLOCSQGVAYLHSMQPKALIHRLDKPPNLLL 164
DB 2031 GGDLLTYLRKARMAATFYGPLLTLDVLDLCVDISKGCYVLERMH---FIHRDLAARNCLV 2087

QY 165 ---VAGGTVLKIGDFTACDI-OTHMTNNKGS---AAMMAPEVPEGSNYSEKCDVFSW 215
DB 2088 SVKDYTSPIRVKIGDFGLARDIYKNDYRKRGEGLLPVRWMAPESLMDGIFTTQSDVNSP 2147

QY 216 GIIWLVIT-RRKPFDEIGGPA---FRIMWAVHNGTR--PPLIKNLKPKIESLMTROCKSW 269
DB 2148 GILIWEILTLGHQPY-----PAHSNLDVLNVYQTGRLEPP--RNCDDLLWNLMTQCWAQ 2200

QY 270 DPSQPSMEEIVKIMTHLMRYFPGADEPLQYPCQYSDGQSNATSSTGSPM----DIAS 325
DB 2201 EPDQRPFTFRIQNQLFRNFFLNS-----IYQCR--DEA--NNSGVINESFEGDGDVICL 2253

QY 326 NTSNKSDDTNMEQVPATNDTIKRLESKLLKNQAKQSQESGRLSLGASHGSSVSELPPTSEG 385
DB 2254 N-----SDDIMPVVLMTKNRGLNVMVLATCGQGEKSEGLGSGQS---ESGCLRKEE 2306

QY 386 KRMSADMSEIEARIAATTNGNQPPRRRSIQDLTVTGTEFG 424
DB 2307 KEPHADKDFCQEKQVAYCPSGKPEGLNYACLTHSGYGDG 2345

Search completed: December 9, 2002, 22:48:21
Job time : 32.1274 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 9, 2002, 18:31:39 ; Search time 60.9474 Seconds
(without alignments)
1265.881 Million cell updates/sec

Title: US-09-830-144-2

Perfect score: 3014

Sequence: 1 MSTAASASSSSSSAGEMIE.....QCKQLEVRSSQKKRQGT 579

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A_Geneseq_101002.*
- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
 - 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
 - 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
 - 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
 - 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
 - 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
 - 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
 - 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
 - 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
 - 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
 - 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
 - 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
 - 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
 - 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
 - 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
 - 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
 - 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
 - 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
 - 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
 - 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
 - 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
 - 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
 - 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3014	100.0	579	18 AAW27093	Human transforming
2	3014	100.0	579	20 AAY09542	Human TAK1 protein
3	3014	100.0	579	21 AAY91000	Human TAK-1 protei
4	3014	100.0	590	20 AAY09547	Human TAK1-6xHis p
5	3006	99.7	579	20 AAY28996	Human TGF-beta act
6	3006	99.7	579	23 ABB85033	Pain regulated pro
7	2986	99.1	579	18 AAW27092	Mouse transforming
8	2982.5	99.0	606	20 AAY28997	Human TGF-beta act
9	2746	91.1	567	20 AAY28998	Human TGF-beta act
10	1310	43.5	261	22 AAU33191	Novel human secret

11	966	32.1	678	22 ABB58061	Drosophila melanog
12	505	16.8	252	22 ABB60985	Drosophila melanog
13	482	16.0	473	22 AAM25322	Human protein sequ
14	481	16.0	455	21 AAB18657	A human regulator
15	481	16.0	455	21 AAY83278	Human survival reg
16	481	16.0	455	21 AAY84321	A human cardiovasc
17	462	15.3	1021	23 ABB61000	Novel human protei
18	461.5	15.3	719	22 AAB85513	Human protein kina
19	461.5	15.3	1036	23 ABB80923	Novel human protei
20	459.5	15.2	847	23 AAE22763	Human mitogen acti
21	456.5	15.1	800	22 AAB71957	Human TGF-beta rec
22	456.5	15.1	800	22 AAB65673	Novel protein kina
23	451.5	15.0	1097	23 AAE21717	Human PKIN-12 prot
24	451	15.0	1046	22 AAE11775	Human kinase (PKIN
25	436.5	14.5	1020	22 ABB58999	Drosophila melanog
26	434.5	14.4	367	21 AAG32053	Arabidopsis thalia
27	434.5	14.4	369	21 AAG32052	Arabidopsis thalia
28	434.5	14.4	407	21 AAG32051	Arabidopsis thalia
29	434	14.4	349	22 AAG75571	Human colon cancer
30	434	14.4	369	21 AAG22172	Arabidopsis thalia
31	434	14.4	374	21 AAG22171	Arabidopsis thalia
32	434	14.4	412	21 AAG22170	Arabidopsis thalia
33	427.5	14.2	341	21 AAG25600	Arabidopsis thalia
34	427.5	14.2	391	21 AAG25599	Arabidopsis thalia
35	426.5	14.2	589	21 AAG45984	Arabidopsis thalia
36	426.5	14.2	732	21 AAG45983	Arabidopsis thalia
37	426.5	14.2	760	21 AAG45982	Arabidopsis thalia
38	424	14.1	977	22 ABB71694	Drosophila melanog
39	418	13.9	835	21 AAB01470	Human CARX (Cardia
40	418	13.9	835	21 AAB65674	Novel protein kina
41	418	13.9	928	22 ABB16533	Novel human diagno
42	414	13.7	835	21 AAB01474	Rat CARX (Cardiac
43	412	13.7	859	16 AAR82886	Human leucine zipp
44	412	13.7	859	18 AAR31227	Human leucine-zipp
45	405	13.4	888	23 ABB57049	Mouse ischaemic co

ALIGNMENTS

RESULT 1
AAW27093
ID AAW27093 standard; Protein; 579 AA.
XX
AC AAW27093;
XX
DT 19-NOV-1997 (first entry)
XX
DE Human transforming growth factor-beta activated kinase TAK-1.
KW TGF-beta; signal transmissio; TGF-beta activated kinase;
KW MAPK kinase activator; AMK-1; bone morphogenetic protein; BMP;
KW protein kinase.
XX
OS Homo sapiens.
XX
PN JP09163990-A.
XX
PD 24-JUN-1997.
XX
PF 27-SEP-1996; 96JP-0256747.
PR 24-JUL-1996; 96US-0685625.
PR 29-SEP-1995; 95JP-0253549.
PA (CHUS) CHUGAI PHARM CO LTD.
PA (UENO/) UENO N.
XX
DR WPI; 1997-380171/35.
DR N-PSDB; AAT85095.
XX
PT DNA encoding transforming growth factor-beta-activated kinase, TAK-1
PT - useful for studying the TGF-beta signal transmission system

XX PS Claim 15; Page 13-15; 20pp; Japanese.

XX CC The present sequence represents human transforming growth factor-beta (TGF-beta) activated kinase, TAK-1. The DNA is used to produce the TAK-1 protein which is involved in the TGF-beta family signal transduction system. TAK-1, also known as activator of MAPK Kinase (AMK-1), is an enzyme which is activated by TGF-beta and bone morphogenetic protein (BMP) and activates MAPK kinase by phosphorylation.

XX SQ Sequence 579 AA;

Query Match 100.0%; Score 3014; DB 18; Length 579;
Best Local Similarity 100.0%; Pred. No. 1.5e-241;
Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTASAAASSSSSAGEMIEAPSOVLNFEIDYKEIEVEEVRGAFGVVCAKAKRAKDV 60
DB 1 MSTASAAASSSSSAGEMIEAPSOVLNFEIDYKEIEVEEVRGAFGVVCAKAKRAKDV 60
QY 61 AIKQIESERKAFIVELRQLSRVNHPIVKLYGACLPVCLVMEYAEGLYVNLHGAE 120
DB 61 AIKQIESERKAFIVELRQLSRVNHPIVKLYGACLPVCLVMEYAEGLYVNLHGAE 120
QY 121 PLPYTTAAHAMSWCLOCSQGVAYLHSMQPKALIHRLDKPNNLLVAGTVLTKICDFTGAC 180
DB 121 PLPYTTAAHAMSWCLOCSQGVAYLHSMQPKALIHRLDKPNNLLVAGTVLTKICDFTGAC 180
QY 181 DIOTHTNNKGSAAWMAPEVEEGSNSEKCDVFSWGIILMEVITRRKPPDEIGGPAFRIM 240
DB 181 DIOTHTNNKGSAAWMAPEVEEGSNSEKCDVFSWGIILMEVITRRKPPDEIGGPAFRIM 240
QY 241 WAVHNGTRPPLIKNLPKPIESLMTRCMSKDPQSPRSMEEIVKIMTHLMRYFPGADEPLQY 300
DB 241 WAVHNGTRPPLIKNLPKPIESLMTRCMSKDPQSPRSMEEIVKIMTHLMRYFPGADEPLQY 300
QY 301 PCQYSDGQSNATSTGSPMDIASTNTSNKSDTNMEQVPATNDTIKRLSKLIKNOAKQ 360
DB 301 PCQYSDGQSNATSTGSPMDIASTNTSNKSDTNMEQVPATNDTIKRLSKLIKNOAKQ 360
QY 361 SESGRSLGASHGSSVESLPPTSEGRKMSADMEIEARIAATTGNGQPRRSIODLTVTG 420
DB 361 SESGRSLGASHGSSVESLPPTSEGRKMSADMEIEARIAATTGNGQPRRSIODLTVTG 420
QY 421 TEBGQVSSRRSSPSVMTTSGPTSEKPTRSHPTPDDSTDTNGSDNSIPMAVLTLDHQL 480
DB 421 TEBGQVSSRRSSPSVMTTSGPTSEKPTRSHPTPDDSTDTNGSDNSIPMAVLTLDHQL 480
QY 481 QPLAPCPNKSMAVFEQHCMAOEYMKVOTETAILLQKQELVAFLLDDEKDOQNTSRL 540
DB 481 QPLAPCPNKSMAVFEQHCMAOEYMKVOTETAILLQKQELVAFLLDDEKDOQNTSRL 540
QY 541 VOEHKKLLDENKSLSTYYQCKKOLEVIRSQOQKRGTS 579
DB 541 VOEHKKLLDENKSLSTYYQCKKOLEVIRSQOQKRGTS 579

PH Key Location/Qualifiers
FT CDS 183..1922
FT /*tag= a

XX MO9921010-A1.
XX **DB 29-APR-1999.**
XX

FF 22-OCT-1998; 98WO-JP04796.
PR 22-OCT-1997; 97JD-0290188.
XX
XX (CHUS) CHUGAI SEIYAKU KK.
XX Ohtomo T, Ono K, Tsuchiya M;
XX
XX WPI; 1999-312645/26.
DR N-PSDB; AAX56279.
XX

PT Screening for TGF- beta inhibitory substances, which are useful as
PT drugs for treatment of diseases relating to its disorder

PS Claim 4; Page 155-157; 195pp; Japanese.

XX CC A method has been developed for screening for substances which inhibit
CC the binding of TAK1 polypeptide to TAK1 polypeptide. The method
CC comprises: (a) contacting the polypeptide in the presence of a sample;
CC and (b) detecting the amount of bound polypeptide, in which the sample
CC can be pre-mixed with TAK1 or TAK1 polypeptide first. The transforming
CC growth factor (TGF)-beta inhibitory substances can be used in drugs for
CC indications e.g. as TGF-beta signal transduction inhibitors or
CC activators, or extracellular matrix protein production enhancement
CC inhibitors or activators, or cell proliferation prevention inhibitors or
CC activators, or monocyte migration inhibitors or activators, or
CC physiological activity induction inhibitors or activators, or
CC immunosuppression inhibitors or activators, or amyloid beta protein
CC precipitation inhibitors or activators, and such substances can also be
CC inhibitors of the TAK1 polypeptide function, particularly kinase
CC activity. The present sequence represents human TAK1.

XX SQ Sequence 579 AA;

Query Match 100.0%; Score 3014; DB 20; Length 579;
Best Local Similarity 100.0%; Pred. No. 1.5e-241;
Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTASAAASSSSSAGEMIEAPSOVLNFEIDYKEIEVEEVRGAFGVVCAKAKRAKDV 60
DB 1 MSTASAAASSSSSAGEMIEAPSOVLNFEIDYKEIEVEEVRGAFGVVCAKAKRAKDV 60
QY 61 AIKQIESERKAFIVELRQLSRVNHPIVKLYGACLPVCLVMEYAEGLYVNLHGAE 120
DB 61 AIKQIESERKAFIVELRQLSRVNHPIVKLYGACLPVCLVMEYAEGLYVNLHGAE 120
QY 121 PLPYTTAAHAMSWCLOCSQGVAYLHSMQPKALIHRLDKPNNLLVAGTVLTKICDFTGAC 180
DB 121 PLPYTTAAHAMSWCLOCSQGVAYLHSMQPKALIHRLDKPNNLLVAGTVLTKICDFTGAC 180
QY 181 DIOTHTNNKGSAAWMAPEVEEGSNSEKCDVFSWGIILMEVITRRKPPDEIGGPAFRIM 240
DB 181 DIOTHTNNKGSAAWMAPEVEEGSNSEKCDVFSWGIILMEVITRRKPPDEIGGPAFRIM 240
QY 241 WAVHNGTRPPLIKNLPKPIESLMTRCMSKDPQSPRSMEEIVKIMTHLMRYFPGADEPLQY 300
DB 241 WAVHNGTRPPLIKNLPKPIESLMTRCMSKDPQSPRSMEEIVKIMTHLMRYFPGADEPLQY 300
QY 301 PCQYSDGQSNATSTGSPMDIASTNTSNKSDTNMEQVPATNDTIKRLSKLIKNOAKQ 360
DB 301 PCQYSDGQSNATSTGSPMDIASTNTSNKSDTNMEQVPATNDTIKRLSKLIKNOAKQ 360
QY 361 SESGRSLGASHGSSVESLPPTSEGRKMSADMEIEARIAATTGNGQPRRSIODLTVTG 420
DB 361 SESGRSLGASHGSSVESLPPTSEGRKMSADMEIEARIAATTGNGQPRRSIODLTVTG 420

QY 421 TEPQVSSRSSPSVRMITTSGPTSEKPTRSHPTPDDSTDTNGSDNSIPMAYLTLDHQL 480
 DB 421 TEPQVSSRSSPSVRMITTSGPTSEKPTRSHPTPDDSTDTNGSDNSIPMAYLTLDHQL 480
 QY 481 OPLAPCNSKESMAVFEQHCMAQYMKVQTEIALLQKQELVAELDDQKQOQNTSRL 540
 DB 481 OPLAPCNSKESMAVFEQHCMAQYMKVQTEIALLQKQELVAELDDQKQOQNTSRL 540
 QY 541 VQEHKLLDENKSLSTYYQCKKQLEVIRSQOQKRGTS 579
 DB 541 VQEHKLLDENKSLSTYYQCKKQLEVIRSQOQKRGTS 579

RESULT 3
 AAY91000
 ID AAY91000 standard; Protein; 579 AA.
 AC AAY91000;
 XX
 DT 04-SEP-2000 (first entry)
 XX
 DE Human TAK-1 protein sequence SEQ ID NO:2.
 XX
 KW Human; TAK-1; TAB-1; mitogen activated protein kinase; MAPK;
 KW screening; signal transduction; inhibition; inflammatory cytokine;
 KW IL-1; interleukin 1; TNF; tumour necrosis factor; inflammation;
 KW antiinflammatory; suppression.
 XX
 OS Homo sapiens.
 XX
 FN WO200023610-A1.
 XX
 PD 27-APR-2000.
 XX
 PF 21-OCT-1999; 99WO-JP05817.
 XX
 PR 21-OCT-1998; 98JP-0299962.
 XX
 PA (CHUS) CHUGAI SEIYAKU KK.
 XX
 PI Tsuchiya M, Ontomo T, Sugamata Y, Matsumoto K;
 DR WPI; 2000-339707/29.
 DR N-PSDB; AAA39105.
 XX
 PT Method for screening inhibitors of TAK1 signal transduction for
 PT suppression of inflammatory cytokine production and use as
 PT antiinflammatory agents -
 XX
 PS Example 1; Page 80-84; 100pp; Japanese.
 XX
 CC The present invention describes a method for screening compounds for
 CC inhibition of inflammatory cytokine signal transduction by contacting
 CC the sample with TAK1 and its receptor TAB1 and selecting for inhibition
 CC of TAK1/TAB1 binding. Also described is a method for screening compounds
 CC for inhibition of inflammatory cytokine signal transduction in which the
 CC inhibition of TAK1 phosphorylation is selected for; and drug
 CC compositions for the treatment of inflammatory disorders containing as
 CC active component an inflammatory cytokine signal transduction inhibitor.
 CC TAK1 is an essential component of the signalling process which results
 CC in release of inflammatory cytokines such as interleukin-1 (IL-1),
 CC IL-10, tumour necrosis factor (TNF) and IL-6. The methods can be used
 CC for the selection of effective antiinflammatory agents. The present
 CC sequence represents human TAK-1, which is used in the exemplification of
 CC the present invention.
 XX
 SQ Sequence 579 AA;

Query Match 100.0%; Score 3014; DB 21; Length 579;
 Best Local Similarity 100.0%; Pred. No. 1.5e-241;
 Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTASAASSSSSSSAGEMIEAPSQVLNFEEIDYKEIEVEEVVGRGAFGVCKAKWRKDV 60
 DB 1 MSTASAASSSSSSSAGEMIEAPSQVLNFEEIDYKEIEVEEVVGRGAFGVCKAKWRKDV 60
 QY 61 AIKQIESSESEKAFIVELRQLSRVNHPIVKLYGACLNPCVLNMEYABGGSLYNVLHGA 120
 DB 61 AIKQIESSESEKAFIVELRQLSRVNHPIVKLYGACLNPCVLNMEYABGGSLYNVLHGA 120
 QY 121 PLPYTYTAHAHMSWCLQCQSGVAYLHSMOPKALIHRLDKPPNLLLVAGTGLVKICDFGTAC 180
 DB 121 PLPYTYTAHAHMSWCLQCQSGVAYLHSMOPKALIHRLDKPPNLLLVAGTGLVKICDFGTAC 180
 QY 181 DIQTHMTNNKGSAAWMAPEVFEKGSYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIM 240
 DB 181 DIQTHMTNNKGSAAWMAPEVFEKGSYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIM 240
 QY 241 WAVHNGTRPPLIKNLPKPIESLMTRCWSKDPSPSMEIEIVKIMTHLMRYFPFGADEPLO 300
 DB 241 WAVHNGTRPPLIKNLPKPIESLMTRCWSKDPSPSMEIEIVKIMTHLMRYFPFGADEPLO 300
 QY 301 PCOYSDGQNSATSGSFMDIASNTSNKSDTNWEOVPATNDTIKRLSKLLKNQAKQQ 360
 DB 301 PCOYSDGQNSATSGSFMDIASNTSNKSDTNWEOVPATNDTIKRLSKLLKNQAKQQ 360
 QY 361 SESGRSLSLGASHGSSVESLPPTSEGRKMSADMSIEARIAAATNGQPRRSIQDLTVTG 420
 DB 361 SESGRSLSLGASHGSSVESLPPTSEGRKMSADMSIEARIAAATNGQPRRSIQDLTVTG 420
 QY 421 TEPQVSSRSSPSVRMITTSGPTSEKPTRSHPTPDDSTDTNGSDNSIPMAYLTLDHQL 480
 DB 421 TEPQVSSRSSPSVRMITTSGPTSEKPTRSHPTPDDSTDTNGSDNSIPMAYLTLDHQL 480
 QY 481 OPLAPCNSKESMAVFEQHCMAQYMKVQTEIALLQKQELVAELDDQKQOQNTSRL 540
 DB 481 OPLAPCNSKESMAVFEQHCMAQYMKVQTEIALLQKQELVAELDDQKQOQNTSRL 540
 QY 541 VQEHKLLDENKSLSTYYQCKKQLEVIRSQOQKRGTS 579
 DB 541 VQEHKLLDENKSLSTYYQCKKQLEVIRSQOQKRGTS 579

RESULT 4
 AAY09547
 ID AAY09547 standard; Protein; 590 AA.
 XX
 AC AAY09547;
 XX
 DT 21-JUL-1999 (first entry)
 XX
 DE Human TAK1-6xHis protein.
 XX
 KW Human; TAB1; TAK1; screening; inhibition; TGF-beta;
 KW transforming growth factor beta.
 XX
 OS Homo sapiens.
 XX Synthetic.
 XX
 PN WO9921010-A1.
 XX
 PD 29-APR-1999.
 XX
 PF 22-OCT-1998; 98WO-JP04796.
 XX
 PR 22-OCT-1997; 97JP-0290188.
 XX
 PA (CHUS) CHUGAI SEIYAKU KK.
 XX
 PI Ontomo T, Ono K, Tsuchiya M;
 XX
 DR WPI; 1999-312645/26.
 DR N-PSDB; AAX56285.
 XX
 PT Screening for TGF- beta inhibitory substances, which are useful as

PT drugs for treatment of diseases relating to its disorder
XX
XX Example 1; Page 171-174; 195pp; Japanese.

CC A method has been developed for screening for substances which inhibit
CC the binding of TAK1 polypeptide to TAK1 polypeptide. The method
CC comprises: (a) contacting the polypeptide in the presence of a sample;
CC and (b) detecting the amount of bound polypeptide, in which the sample
CC can be pre-mixed with TAK1 or TAK1 polypeptide fixat. The transforming
CC growth factor (TGF)-beta inhibitory substances can be used in drugs for
CC indications e.g. as TGF-beta signal transduction inhibitors or
CC activators, or extracellular matrix protein production enhancement
CC inhibitors, or activators, or cell proliferation prevention inhibitors or
CC activators, or monocyte migration inhibitors or activators, or
CC physiological activity induction inhibitors or activators, or
CC immunosuppression inhibitors or activators, or amyloid beta protein
CC precipitation inhibitors or activators, and such substances can also be
CC inhibitors of the TAK1 polypeptide function, particularly kinase
CC activity. The present sequence represents TAK1-6xHis from an example of
CC the present invention.
XX
XX

SQ Sequence 590 AA;

Query Match 100.0%; Score 3014; DB 20; Length 590;
Best Local Similarity 100.0%; Pred. No. 1.5e-241;
Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTASASSSSSSAGEMIEAPSOVLNFEIDYKEIEVEEVVGRGAFVCAKMRKADY 60
DB 1 MSTASASSSSSSAGEMIEAPSOVLNFEIDYKEIEVEEVVGRGAFVCAKMRKADY 60
QY AIKQIESSEKKAFFVELRQLSRVNHPNIVKLYGACLNPCVMEVAGGSLYNVLHGAE 120
DB AIKQIESSEKKAFFVELRQLSRVNHPNIVKLYGACLNPCVMEVAGGSLYNVLHGAE 120
QY 121 PLPYTTAAHAMSWCLQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGTVLTKICDFGTAC 180
DB 121 PLPYTTAAHAMSWCLQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGTVLTKICDFGTAC 180
QY 181 DIQHTMTNNKSAAMMAPEVFEESGNSSEKCDVFSWGIILMEVITRRKPPDEIGGPAFRIM 240
DB 181 DIQHTMTNNKSAAMMAPEVFEESGNSSEKCDVFSWGIILMEVITRRKPPDEIGGPAFRIM 240
QY 241 WAVHNGTRPPLIKNLPKPIESLMTRCMSKDPQSPMEIEIVKIMTHLMRYFPGADEPLQY 300
DB 241 WAVHNGTRPPLIKNLPKPIESLMTRCMSKDPQSPMEIEIVKIMTHLMRYFPGADEPLQY 300
QY 301 PCQYSDEGQNSATSTGSPMDIASTNTSNKSDTNMEQVPATNDTIKRLSKLLKNQAKQ 360
DB 301 PCQYSDEGQNSATSTGSPMDIASTNTSNKSDTNMEQVPATNDTIKRLSKLLKNQAKQ 360
QY 361 SEGGRLSLGASHGSSVESLPTSEGRKMSADMEIEARIAATGNGQPRRSIQDITLVG 420
DB 361 SEGGRLSLGASHGSSVESLPTSEGRKMSADMEIEARIAATGNGQPRRSIQDITLVG 420
QY 421 TEBGOVSSRSRSPSVNRMTTSGPTSEKPTRSHPTPDSDTNGSDNSIPMAYLTLDHOL 480
DB 421 TEBGOVSSRSRSPSVNRMTTSGPTSEKPTRSHPTPDSDTNGSDNSIPMAYLTLDHOL 480
QY 481 QPLAPCNSKESMAVFEQHKMAQEVYKQVTEITALLQKQELVALLDQEKQNTSRL 540
DB 481 QPLAPCNSKESMAVFEQHKMAQEVYKQVTEITALLQKQELVALLDQEKQNTSRL 540
QY 541 VOEHKKLLDENKSLSTYYQCKKQLEVIIRSQOQKROGTS 579
DB 541 VOEHKKLLDENKSLSTYYQCKKQLEVIIRSQOQKROGTS 579

RESULT 5
AA288996
ID AA288996 standard; Protein; 579 AA.
XX
AC AA288996;

XX 29-OCT-1999 (first entry)
DT
XX
XX Human TGF-beta activated kinase (TAK) 1a amino acid sequence.
DE
XX

KW Nuclear factor kappa B; NF-KB; inhibitor; TGF-beta activated kinase 1;
KW TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;
KW intractable disease; atrophic dermatitis; psoriasis; viral infection;
KW endotoxin shock; septicemia; human; hTAK1a.
XX

OS Homo sapiens.

PN MO3940202-A1.

PD 12-AUG-1999.

PF 02-FEB-1999; 99MO-JP00422.

PR 30-OCT-1998; 98JP-0309316.

PR 06-FEB-1998; 98JP-0026003.

PA (TANA) TANABE SEIYAKU CO.

PI Hasegawa K, Kageyama N, Sakurai H, Sugita T;

DR WP1; 1999-494298/41.

DR N-PSDB; AAX99696.

PT Nuclear factor kappa B activation inhibitors, useful as preventives

for, e.g. autoimmune diseases

Examples; Page 35-39; 49pp; Japanese.

CC The invention provides a method for identifying or screening a nuclear
CC factor kappa B (NF-kB) activation inhibitor by examining the effect of a
CC test substance on modulating the function(s) of TGF-beta activated kinase
CC 1 (TAK1). The NFkB activation inhibitors targeting on TAK1 can be used to
CC treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis),
CC intractable diseases with inflammation (such as atrophic dermatitis and
CC psoriasis), viral infection, endotoxin shock, septicemia and others. The
CC present sequence represents the amino acid sequence of human TAK1a
CC (hTAK1a) protein.
XX
XX

SQ Sequence 579 AA;

Query Match 99.7%; Score 3006; DB 20; Length 579;
Best Local Similarity 99.8%; Pred. No. 6.8e-241;
Matches 576; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTASASSSSSSAGEMIEAPSOVLNFEIDYKEIEVEEVVGRGAFVCAKMRKADY 60
DB 1 MSTASASSSSSSAGEMIEAPSOVLNFEIDYKEIEVEEVVGRGAFVCAKMRKADY 60
QY 61 AIKQIESSEKKAFFVELRQLSRVNHPNIVKLYGACLNPCVMEVAGGSLYNVLHGAE 120
DB 61 AIKQIESSEKKAFFVELRQLSRVNHPNIVKLYGACLNPCVMEVAGGSLYNVLHGAE 120
QY 121 PLPYTTAAHAMSWCLQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGTVLTKICDFGTAC 180
DB 121 PLPYTTAAHAMSWCLQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGTVLTKICDFGTAC 180
QY 181 DIQHTMTNNKSAAMMAPEVFEESGNSSEKCDVFSWGIILMEVITRRKPPDEIGGPAFRIM 240
DB 181 DIQHTMTNNKSAAMMAPEVFEESGNSSEKCDVFSWGIILMEVITRRKPPDEIGGPAFRIM 240
QY 241 WAVHNGTRPPLIKNLPKPIESLMTRCMSKDPQSPMEIEIVKIMTHLMRYFPGADEPLQY 300
DB 241 WAVHNGTRPPLIKNLPKPIESLMTRCMSKDPQSPMEIEIVKIMTHLMRYFPGADEPLQY 300
QY 301 PCQYSDEGQNSATSTGSPMDIASTNTSNKSDTNMEQVPATNDTIKRLSKLLKNQAKQ 360
DB 301 PCQYSDEGQNSATSTGSPMDIASTNTSNKSDTNMEQVPATNDTIKRLSKLLKNQAKQ 360

QY 361 SESGRSLGASHGSSVESLPTSEGRKMSADMSIEARIAATTGNGOPRRRSIQDLTVTG 420
 DB 361 SESGRSLGASHGSSVESLPTSEGRKMSADMSIEARIAATTGNGOPRRRSIQDLTVTG 420
 QY 421 TEPQVSSRSSPSVRMITTSGPTSEKPTRSHPTPDDSTDITNGSDNSIPMAYLTLDHQL 480
 DB 421 TEPQVSSRSSPSVRMITTSGPTSEKPTRSHPTPDDSTDITNGSDNSIPMAYLTLDHQL 480
 QY 481 QPLAPCNSKESMAVFEQHCCKMAQVYMKVOTETALLLQKQELVAELDQDKQQTNSRL 540
 DB 481 QPLAPCNSKESMAVFEQHCCKMAQVYMKVOTETALLLQKQELVAELDQDKQQTNSRL 540
 QY 541 VQEHKLLDENKSLSTYYQCKKQLEVIRSQQQKROGTS 579
 DB 541 VQEHKLLDENKSLSTYYQCKKQLEVIRSQQQKROGTS 579
 RESULT 6
 ABB85033
 ID ABB85033 standard; Protein; 579 AA.
 XX AC ABB85033;
 XX DT 16-MAY-2002 (first entry)
 XX Pain regulated protein sequence 28.
 XX Pain; analgesic; gene therapy; neurological disorder;
 KW neurodegenerative disease.
 XX OS Homo sapiens.
 XX PN WO200212338-A2.
 XX PD 14-FEB-2002.
 XX PF 03-AUG-2001; 2001WO-EP09011.
 XX PR 03-AUG-2000; 2000DE-1037759.
 XX PA (CHEF) GRUENENTHAL GMBH.
 XX PI Gillen C, Wetzels I, Wnendt S, Weihe E, Schaefer MK;
 XX WPI; 2002-257469/30.
 DR N-PSDB; ABL88437.
 XX Identifying pain-regulating compounds, useful for treating chronic pain
 PT and for diagnosis, by measuring binding of compounds to specific
 PT peptides and proteins
 XX Claim 1; Fig 44; 213pp; German.
 XX The invention relates to identifying pain-regulating substances (A)
 CC comprises (i) incubating a test substance with a cell (or preparation
 CC from it) that has synthesised a peptide or protein (B) and (ii) measuring
 CC either binding of the test substance to (B) or some functional parameter
 CC that is altered by this binding. The method is useful for identifying
 CC pain-regulating substances (A) with analgesic activity. (A) along with
 CC nucleic acid (ABL88411-ABL88441) that encode proteins (B)
 CC ABB85006-ABB85037) that interact with (A); (B); vectors containing the
 CC nucleic acid; antibodies against (B); cells that express (B) and agents
 CC that bind to (B), are all useful for treating pain, particularly chronic
 CC pain, including use in gene therapy. The same materials can also be used
 CC for diagnosis, e.g. of neurological and neurodegenerative diseases. The
 CC present sequence is that of a polypeptide of the invention.
 XX SQ Sequence 579 AA;
 Query Match 99.7%; Score 3006; DB 23; Length 579;
 Best Local Similarity 99.8%; Pred. No. 6.8e-241;
 Matches 578; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 METASAAASSSSSSAGEMIEAPSOVLNFEEDYKBEIEVEVVGAFGVCAKWKRAKDV 60
 DB 1 METASAAASSSSSSAGEMIEAPSOVLNFEEDYKBEIEVEVVGAFGVCAKWKRAKDV 60
 QY 61 AIKQIESSESEKAFIVELRQLSRVNHPNIVKLYGACLNVPCLVMYEAEGGSLYNNVLHGA 120
 DB 61 AIKQIESSESEKAFIVELRQLSRVNHPNIVKLYGACLNVPCLVMYEAEGGSLYNNVLHGA 120
 QY 121 PLPYYTAAAHMSWCLQCSQGVAYLHSMQPKALIHRLDKPPNLLLVAGGTVLKICDFGTAC 180
 DB 121 PLPYYTAAAHMSWCLQCSQGVAYLHSMQPKALIHRLDKPPNLLLVAGGTVLKICDFGTAC 180
 QY 181 DIQHTMTNNKGSAAWMADEVFEGSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIM 240
 DB 181 DIQHTMTNNKGSAAWMADEVFEGSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIM 240
 QY 241 WAVHNGTRPPLIKNLKPPIESLMTTRCWSKDPSPSMEIIVKIMTHLMRYEPFGADEPLQY 300
 DB 241 WAVHNGTRPPLIKNLKPPIESLMTTRCWSKDPSPSMEIIVKIMTHLMRYEPFGADEPLQY 300
 QY 301 PCQYSDGQSNASATSGSFMDIASNTNSKSDTNNEOVPAATNDTIKRLSKLLKNQAKQQ 360
 DB 301 PCQYSDGQSNASATSGSFMDIASNTNSKSDTNNEOVPAATNDTIKRLSKLLKNQAKQQ 360
 QY 361 SESGRSLGASHGSSVESLPTSEGRKMSADMSIEARIAATTGNGOPRRRSIQDLTVTG 420
 DB 361 SESGRSLGASHGSSVESLPTSEGRKMSADMSIEARIAATTGNGOPRRRSIQDLTVTG 420
 QY 421 TEPQVSSRSSPSVRMITTSGPTSEKPTRSHPTPDDSTDITNGSDNSIPMAYLTLDHQL 480
 DB 421 TEPQVSSRSSPSVRMITTSGPTSEKPTRSHPTPDDSTDITNGSDNSIPMAYLTLDHQL 480
 QY 481 QPLAPCNSKESMAVFEQHCCKMAQVYMKVOTEIALLQKQELVAELDQDKQQTNSRL 540
 DB 481 QPLAPCNSKESMAVFEQHCCKMAQVYMKVOTEIALLQKQELVAELDQDKQQTNSRL 540
 QY 541 VQEHKLLDENKSLSTYYQCKKQLEVIRSQQQKROGTS 579
 DB 541 VQEHKLLDENKSLSTYYQCKKQLEVIRSQQQKROGTS 579
 RESULT 7
 AAW27092
 ID AAW27092 standard; Protein; 579 AA.
 XX AC AAW27092;
 XX DT 19-NOV-1997 (first entry)
 XX Mouse transforming growth factor-beta activated kinase TAK-1.
 DE TGF-beta; signal transduction; TGF-beta activated kinase;
 KW MAPK kinase activator; AMK-1; bone morphogenetic protein; BMP;
 KW protein kinase.
 XX Mus musculus.
 OS JP09163990-A.
 XX 24-JUN-1997.
 XX 27-SEP-1996; 96JP-0256747.
 XX 29-JUL-1996; 96US-0685625.
 PR 24-SEP-1995; 95JP-0253549.
 XX (CHUS) CHUGAI PHARM CO LTD.
 PA (UENO/) UENO N.
 XX WPI; 1997-380171/35.
 DR N-PSDB; AAT85094.
 XX DNA encoding transforming growth factor-beta-activated kinase, TAK-1

PT - useful for studying the TGF-beta signal transmission system
 XX
 PS Claim 14; Page 10-12; 20pp; Japanese.

XX The present sequence represents mouse transforming growth factor-beta
 CC (TGF-beta) activated kinase, TAK-1. The DNA is used to produce the
 CC TAK-1 protein which is involved in the TGF-beta family signal
 CC transmission system. TAK-1, also known as activator of MAPK kinase
 CC (AMK-1), is an enzyme which is activated by TGF-beta and bone
 CC morphogenetic protein (BMP) and activates MAPK kinase by
 CC phosphorylation.

XX Sequence 579 AA;

Query Match 99.1%; Score 2986; DB 18; Length 579;
 Best Local Similarity 99.1%; Pred. No. 3,1e-239;
 Matches 574; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSTASASSSSSSAGEMIAPSOVLNFEEDYKEIEVEEVGGAFGVYCKAKMRADY 60
 DB 1 MSTASASSSSSSSASEMIAPSOVLNFEEDYKEIEVEEVGGAFGVYCKAKMRADY 60
 QY AIKQIESESEKAFIVELRQLSRVNHPIVLYGACLNPCVLMVEYAGGSLYNVLHGAE 120
 DB AIKQIESESEKAFIVELRQLSRVNHPIVLYGACLNPCVLMVEYAGGSLYNVLHGAE 120
 QY PLPYTTAAHMSWCLQCSQGVAYLHSMQPKALIHDLKPNLLLVAGGTVLKICDFGTAC 180
 DB PLPYTTAAHMSWCLQCSQGVAYLHSMQPKALIHDLKPNLLLVAGGTVLKICDFGTAC 180
 QY 181 DIQHTMTNNKGSAAWMAPEVEGGSNYSEKCDVFSWGIILMEVITRRKPFDEIGGPAFRIM 240
 DB 181 DIQHTMTNNKGSAAWMAPEVEGGSNYSEKCDVFSWGIILMEVITRRKPFDEIGGPAFRIM 240
 QY 241 WAVHNGTRPPLIKNLKPIESLMTRCWSDPQSPSEMEIVKIMTHLMRYFPGADEPLQY 300
 DB 241 WAVHNGTRPPLIKNLKPIESLMTRCWSDPQSPSEMEIVKIMTHLMRYFPGADEPLQY 300
 QY 301 PCOYSDGQNSATSTGSMFMDIASTNTSNKSDTNMEQVPATNDTIKLESKLLKNQAKQ 360
 DB 301 PCOYSDGQNSATSTGSMFMDIASTNTSNKSDTNMEQVPATNDTIKLESKLLKNQAKQ 360
 QY 361 SESGRSLGASHGSSVESLPTSEGRKMSADNSEIARIAATTTGNCQPRRSIODLTVTG 420
 DB 361 SESGRSLGASHGSSVESLPTSEGRKMSADNSEIARIAATTTGNCQPRRSIODLTVTG 420
 QY 421 TEEGQVSSRRSSPSVMTTSSGPTSEKPRSHWPTPDDSTDTNGSDNSIPMAVLTLDHQL 480
 DB 421 TEEGQVSSRRSSPSVMTTSSGPTSEKPRSHWPTPDDSTDTNGSDNSIPMAVLTLDHQL 480
 QY 481 QPLAPCPNSKESMAVEHQKMAQEVYKQTEITALLLQKQELVAELDDEKDOQNTSRL 540
 DB 481 QPLAPCPNSKESMAVEHQKMAQEVYKQTEITALLLQKQELVAELDDEKDOQNTSRL 540
 QY 541 VQEHKKLLDENKSLSTYYQCKKQLEVINSOQKRGQTS 579
 DB 541 VQEHKKLLDENKSLSTYYQCKKQLEVINSOQKRGQTS 579

RESULT 8
 AAAY28997
 ID AAAY28997 standard; Protein; 606 AA.

XX AAAY28997;

XX 29-OCT-1999 (first entry)

XX Human TGF-beta activated kinase (TAK) 1b amino acid sequence.

XX Nuclear factor kappa B; NF-KB; inhibitor; TGF-beta activated kinase 1;
 KW TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;
 KW intractable disease; atrophic dermatitis; psoriasis; viral infection;
 KW endotoxin shock; septicemia; human; htk1b.

XX OS Homo sapiens.

XX PN W09940202-A1.

XX PD 12-AUG-1999.

XX PF 02-FEB-1999; 99WO-JP00422.

XX PR 30-OCT-1998; 98UP-0309316.

XX PR 06-FEB-1998; 98UP-0026003.

XX (TANA) TANABE SEIYAKU CO.

XX PI Hasegawa K, Kageyama N, Sakurai H, Sugita T;

XX WPI; 1999-494298/41.

XX DR N-PADB; AAX9697.

XX PS Examples; Page 39-43; 49pp; Japanese.

CC The invention provides a method for identifying or screening a nuclear
 CC factor kappa B (NF-kB) activation inhibitor by examining the effect of a
 CC test substance on modulating the function(s) of TGF-beta activated kinase
 CC 1 (TAK1). The NFkB activation inhibitors targeting on TAK1 can be used to
 CC treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis),
 CC intractable diseases with inflammation (such as atrophic dermatitis and
 CC psoriasis), viral infection, endotoxin shock, septicemia and others. The
 CC present sequence represents the amino acid sequence of human TAK1b
 CC (htk1b) protein.

XX SQ Sequence 606 AA;

Query Match 99.0%; Score 2982.5; DB 20; Length 606;
 Best Local Similarity 95.4%; Pred. No. 6,6e-239;
 Matches 578; Conservative 0; Mismatches 1; Indels 27; Gaps 1;

QY 1 MSTASASSSSSSAGEMIAPSOVLNFEEDYKEIEVEEVGGAFGVYCKAKMRADY 60
 DB 1 MSTASASSSSSSSAGEMIAPSOVLNFEEDYKEIEVEEVGGAFGVYCKAKMRADY 60
 QY AIKQIESESEKAFIVELRQLSRVNHPIVLYGACLNPCVLMVEYAGGSLYNVLHGAE 120
 DB AIKQIESESEKAFIVELRQLSRVNHPIVLYGACLNPCVLMVEYAGGSLYNVLHGAE 120
 QY 61 AIKQIESESEKAFIVELRQLSRVNHPIVLYGACLNPCVLMVEYAGGSLYNVLHGAE 180
 DB 61 AIKQIESESEKAFIVELRQLSRVNHPIVLYGACLNPCVLMVEYAGGSLYNVLHGAE 180
 QY 181 DIQHTMTNNKGSAAWMAPEVEGGSNYSEKCDVFSWGIILMEVITRRKPFDEIGGPAFRIM 240
 DB 181 DIQHTMTNNKGSAAWMAPEVEGGSNYSEKCDVFSWGIILMEVITRRKPFDEIGGPAFRIM 240
 QY 241 WAVHNGTRPPLIKNLKPIESLMTRCWSDPQSPSEMEIVKIMTHLMRYFPGADEPLQY 300
 DB 241 WAVHNGTRPPLIKNLKPIESLMTRCWSDPQSPSEMEIVKIMTHLMRYFPGADEPLQY 300
 QY 301 PCOYSDGQNSATSTGSMFMDIASTNTSNKSDTNMEQVPATNDTIKLESKLLKNQAKQ 360
 DB 301 PCOYSDGQNSATSTGSMFMDIASTNTSNKSDTNMEQVPATNDTIKLESKLLKNQAKQ 360
 QY 361 SESGRSLGASHGSSVESLPTSEGRKMSADNSEIARIAATTTGNCQPRRSIODLTVTG 420
 DB 361 SESGRSLGASHGSSVESLPTSEGRKMSADNSEIARIAATTTGNCQPRRSIODLTVTG 420
 QY 421 ILDPVEIYISNGQPRRSIODLTVTGTEPQVSSRRSSPSVMTTSSGPTSEKPRSHWPTPDDSTDTNGSDNSIPMAVLTLDHQL 480
 DB 421 ILDPVEIYISNGQPRRSIODLTVTGTEPQVSSRRSSPSVMTTSSGPTSEKPRSHWPTPDDSTDTNGSDNSIPMAVLTLDHQL 480
 QY 481 WTPDSTDTNGSDNSIPMAVLTLDHQLQPLAPCPNSKESMAVEHQKMAQEVYKQTEI 513

```

Db 481 WTPDDSTDTNGSDNSIPMAYLTLDHQLQPLAPCPNSKESMAVFEQCKMAQYMKVQTEI 540
QY 514 ALLLQKQELVAELDDEKQDQNTSRLVQEHKLLDENKSLSTYYQCKKQLEVIRSQQQ 573
Db 541 ALLLQKQELVAELDQDEKQDQNTSRLVQEHKLLDENKSLSTYYQCKKQLEVIRSQQQ 600
QY 574 KRQGTS 579
Db 601 KRQGTS 606

RESULT 9
AAV28998
ID AAY28998 standard; Protein; 567 AA.
XX
AC AAY28998;
XX
DT 29-OCT-1999 (first entry)
XX
DE Human TGF-beta activated kinase (TAK) 1c amino acid sequence.
XX
KW Nuclear factor kappa B; NF-kB; inhibitor; TGF-beta activated kinase 1;
KW TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;
KW intractable disease; atrophic dermatitis; psoriasis; viral infection;
KW endotoxin shock; septicemia; human; hTAK1c.
XX
OS Homo sapiens.
XX
PN WO9940202-A1.
XX
PD 12-AUG-1999.
XX
PF 02-FEB-1999; 99WO-JF00422.
XX
PR 30-OCT-1998; 98JP-0309316.
PR 06-FEB-1998; 98JP-0026003.
XX
(TANA ) TANABE SEYAKU CO.
XX
Hasegawa K, Kageyama N, Sakurai H, Sugita T;
WPI; 1999-494298/41.
DR N-PSDB; AAX99698.
XX
Nuclear factor kappa B activation inhibitors, useful as preventives
for, e.g. autoimmune diseases
XX
Examples; Page 43-46; 49pp; Japanese.
XX
The invention provides a method for identifying or screening a nuclear
factor kappa B (NF-kB) activation inhibitor by examining the effect of a
test substance on modulating the function(s) of TGF-beta activated kinase
1 (TAK1). The NFkB activation inhibitors targeting on TAK1 can be used to
treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis),
intractable diseases with inflammation (such as atrophic dermatitis and
psoriasis), viral infection, endotoxin shock, septicemia and others. The
present sequence represents the amino acid sequence of human TAK1c
(hTAK1c) protein.
XX
SQ Sequence 567 AA;

Query Match 91.1%; Score 2746; DB 20; Length 567;
Best Local Similarity 88.8%; Pred. No. 2.7e-219;
Matches 538; Conservative 0; Mismatches 2; Indels 66; Gaps 2;

QY 1 MSTASAASSSSSSAGMIEAPQVLNFEIDYKEIEVEEVGGRGAFGVVCKAKWRAKDV 60
Db 1 MSTASAASSSSSSAGMIEAPQVLNFEIDYKEIEVEEVGGRGAFGVVCKAKWRAKDV 60
QY 61 AIKQIESESRKAFIVELRQLSRVNHPIVKLYGACLNVPCLVMEYAEAGSLYNVLHGAE 120
Db 61 AIKQIESESRKAFIVELRQLSRVNHPIVKLYGACLNVPCLVMEYAEAGSLYNVLHGAE 120

```

```

QY 121 PLPYTTAAHAMSWCLQCSQGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFGTAC 180
Db 121 PLPYTTAAHAMSWCLQCSQGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFGTAC 180
QY 181 DIQTHMTNNKGSAAWMAPEVPEGSGNYSEKCDVFSWGIIILWEVITRRKPPFDEIGGPAFRIM 240
Db 181 DIQTHMTNNKGSAAWMAPEVPEGSGNYSEKCDVFSWGIIILWEVITRRKPPFDEIGGPAFRIM 240
QY 241 WAVHNGTRPPLIKNLKPISLMTRCWSKQSPRSMEIEIVKIMTHLMRYFPFGADEPLQY 300
Db 241 WAVHNGTRPPLIKNLKPISLMTRCWSKQSPRSMEIEIVKIMTHLMRYFPFGADEPLQY 300
QY 301 PCQYSDEGQSN SATSGSFMDIASTNTSNKSDTNMEQVPATNDTIKRLESKLLKNQAKQQ 360
Db 301 PCQYSDEGQSN SATSGSFMDIASTNTSNKSDTNMEQVPATNDTIKRLESKLLKNQAKQQ 360
QY 361 SESGRSLGASHGSSVESLPTTSEGKMSADMSEIARIAATT----- 403
Db 361 SESGRSLGASHGSSVESLPTTSEGKMSADMSEIARIAATTAYSXPKRGRHKRTASFGN 420
QY 404 -----GNGOPRRRSIQDLTVTGTGPGQVSSRSSSPSVRMITTSGPTSEKPTRSHP 453
Db 421 ILDPVEIVISGNGOPRRRSIQDLTVTGTGPGQVSSRSSSPSVRMITTSGPTSEKPTRSHP 480
QY 454 WTPDDSTDTNGSDNSIPMAYLTLDHQLQPLAPCPNSKESMAVFEQCKMAQYMKVQTEI 513
Db 481 WTPDDSTDTNGSDNSIPMAYLTLDHQLQ----- 508
QY 514 ALLLQKQELVAELDQDEKQDQNTSRLVQEHKLLDENKSLSTYYQCKKQLEVIRSQQQ 573
Db 509 -----QELVAELDQDEKQDQNTSRLVQEHKLLDENKSLSTYYQCKKQLEVIRSQQQ 561
QY 574 KRQGTS 579
Db 562 KRQGTS 567

RESULT 10
AAU33191
ID AAU33191 standard; Protein; 261 AA.
XX
AC AAU33191;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human secreted protein #3682.
XX
KW Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200179449-A2.
XX
PD 25-OCT-2001.
XX
PF 16-APR-2001; 2001WO-US08656.
XX
PR 18-APR-2000; 2000US-0552929.
PR 26-JAN-2001; 2001US-0770160.
XX
(HYSE-) HYSEQ INC.
XX
Tang YT, Liu C, Drmanac RT;
XX
WPI; 2001-611725/70.
XX
Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy -
XX
Claim 20; Page 726; 765pp; English.

```

XX The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising
CC the nucleic acids encoding the polypeptides and cells genetically
CC engineered to express them are also useful for producing the proteins.
CC The proteins are useful in genetic vaccination, testing and
CC therapy, and can be used as nutritional supplements. They may be used to
CC increase stem cell proliferation; to regulate haematopoiesis; and in
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
CC immune suppression and/or stimulation; as anti-inflammatory agents; and
CC in treatment of leukemias. AAU29510-AAU3304 represent the amino acid
CC sequences of novel human secreted proteins of the invention.

XX Sequence 261 AA;

Query Match 43.5%; Score 1310; DB 22; Length 261;

Best Local Similarity 99.6%; Pred. No. 1.8e-100; Mismatches 1; Indels 0; Gaps 0;

Matches 259; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DB 1 MDIASTNTSNKSDPTNMEQVPAINDTIKRLSKLLKNQAQOSESGLSIGASHGSSVESL 60

QY 380 PPTSEGRKMSADMSEIEARIAATTNGQPPRRSIODLTVTGTEPGVSSRSSPSVRMT 439

DB 61 PPTSEGRKMSADMSEIEARIAATTNGQPPRRSIODLTVTGTEPGVSSRSSPSVRMT 120

QY 440 TSPPTSEKPRSHPMPTPDDSTDNSSDINSIPMAVLTLDHQLQPLACPSKESMAVFEQH 499

DB 121 TSGPTSEKPRSHPMPTPDDSTDNSSDINSIPMAVLTLDHQLQPLACPSKESMAVFEQH 180

QY 500 CKMAQEMKVQTEIALLQKQELVAELDDKQDQNTSRLVQEHKKLLDENKSLSTYYQ 559

DB 181 CKMAQEMKVQTEIALLQKQELVAELDDKQDQNTSRLVQEHKKLLDENKSLSTYYQ 240

QY 560 QCKKQLEVTSSQOQKRGTS 579

DB 241 QCKKQLEVTSSQOQKRGTS 260

RESULT 11

ABBS8061 ID ABBS8061 standard; Protein; 678 AA.

XX ABBS8061;

DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 975.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PMD, Myers EW;

XX WPI; 2001-656860/75.
DR N-PSDB; ABL02164.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX Disclosure; SEQ ID NO 975; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 678 AA;

Query Match 32.1%; Score 966; DB 22; Length 678;

Best Local Similarity 36.0%; Pred. No. 3e-71; Mismatches 207; Indels 98; Gaps 14;

Matches 226; Conservative 97; Mismatches 207; Indels 98; Gaps 14;

QY 18 MIEAPQVLNPEEIDYKEIEVEEVYVGRGAFQVCKAKWRAKDVAIKQIESSEKKAFFIVE 77

DB 1 MATSLDALQAAVDFSEBITTRKRVGSGSYGVCKAVMRKLAIVAKVFFSAQKDIKE 60

QY 78 LROLSRVNHPNIVKLYG-ACLNPFVCLMEYABGSDLYNVLHGAEPPLPYTAHANMSWCL 135

DB 61 VKQLSRVNHPNIIALHGSIYQATYLLIMEPAGGSLHNFLLG-KVPAVYSLHAAMSWAR 119

QY 136 QCSQGVATLHSMQKALIRHDLKPPNLLVAGGVTLKICPFGNACDIQTMTNNKGSAAV 195

DB 120 QCAAGLALYLAHMTPKPLIHRDVKPLNLLTNKGNLKICDFGVADASTMTNNRGSAAV 179

QY 196 MAPEVFEBSNYSKCDVFSNGIILMEVITRRKPEDEGSAFRIMAVHNGTSPPLIKNL 255

DB 180 MAPEVFEBSKYTEKCDIFSMALIVMEVLVSKQPKGIDN-AVITQMKIYAGEPPLITTC 238

QY 256 PKPIESLMTRCXKDPGQPSMEBIVKIMTHLMKRYFPGADEPLYOY----- 300

DB 239 PKRIEDLMTACWKTVPEDRPSMGYIVGMHEIVYDYGADKALEYTFVNGQIVTKESDGT 298

QY 301 -----PCQYSDQGSN-----SATSTGSPMDIASTNTSNKSDTN 334

DB 299 VAAQPDLSLSQEGELSPSSQTLPPTAANNANVNAIAISKTTTSMTEINTSSTDITPTN 358

QY 335 MEQVP-----ATN--DTIKRLSKLLKN-----QAKQOSESGRSLGASHSSV 376

DB 359 SGQLDNNPFLPYMTNRWDALPEEBSNESRNDSEFNLSAATQRLLETIRNGMLMACKPM 418

QY 377 ESLPPTSEKMSADMEIEARIAATTNGQPPRRSIODLTVTGTGP-----GQ 425

DB 419 EQL--TLDVANGFDLSPSSSSSTNAKSDGR-----LTVIDTKVMMTTTDLSSNNGG 472

QY 426 VSSRSSSPSVAMITTSGETSEKPTRSHPW-----TPDDSTDTNGSNSIPMAVLT 476

DB 473 IHAHSNGILSHANGWOARDELOQEHQEIIVSLDVDPDDEENDGTQSLAE---IL 529

QY 477 DHOLOPLAPCPSKESMAVFEQHQCKMAQEMKVQTEIALLQKQELVAELDDKQDQON 536

DB 530 DPLOPEPPIPNDAESQULYRDRHMAKEYLSVDTNLYVADQFDKULVOMDTERQK- 588

QY 537 TSLRVOEHKKLLDENKSLSTYYQCKKQ 564

DB 589 -----QELLRKMKDKKQGLQSLYNNLQOQ 611


```

RESULT 12
ABB60985
ID   ABB60985 standard; Protein; 252 AA.
XX   AC
XX   ABB60985;
XX   DT
XX   26-MAR-2002 (first entry)
XX   DE
XX   Drosophila melanogaster polypeptide SEQ ID NO 9747.
XX   DE
XX   Drosophila; developmental biology; cell signalling; insecticide;
XX   KW
XX   pharmaceutical.
XX   OS
XX   Drosophila melanogaster.
XX   PN
XX   WO200171042-A2.
XX   PD
XX   27-SEP-2001.
XX   PF
XX   23-MAR-2001; 2001WO-US09231.
XX   PR
XX   23-MAR-2000; 2000US-191637P.
XX   PR
XX   11-JUL-2000; 2000US-0614150.
XX   PA
XX   (PEXE ) PE CORP NY.
XX   PI
XX   Venter JC, Adams M, Li PWD, Myers EW;
XX   WPI; 2001-656860/75.
XX   DR
XX   N-PSDB; ABL05088.
XX   PT
XX   New isolated nucleic acid detection reagent for detecting 1000 or more
XX   genes from Drosophila and for elucidating cell signalling and cell-cell
XX   interactions -
XX   PS
XX   Disclosure; SEQ ID NO 9747; 21bp + Sequence Listing; English.
XX   CC
XX   The invention relates to an isolated nucleic acid detection reagent
XX   capable of detecting 1000 or more genes from Drosophila. The invention is
XX   useful in developmental biology and in elucidating cell signalling and
XX   cell-cell interactions in higher eukaryotes for the development of
XX   insecticides, therapeutics and pharmaceutical drugs. The invention
XX   discloses genomic DNA sequences (ABL01840-ABL16175), expressed DNA
XX   sequences (ABL01840-ABL16175) and the encoded proteins
XX   (ABBS7737-ABBS72072).
XX   CC
XX   The sequence data for this patent did not form part of the printed
XX   specification, but was obtained in electronic format directly from WIPO
XX   at ftp.wipo.int/pub/published_pct_sequences.
XX   SQ
XX   Sequence 252 AA;

Query Match
Best Local Similarity 16.8%; Score 505; DB 22; Length 252;
Matches 99; Conservative 48; Mismatches 71; Indels 26; Gaps 4;

QY 29 EEIDYKEIEVEYVGARGVCKAKRQDVAKIOIESERKAPFIVELRLQSLRVNHPN 88
Db 6 EGVPEEIIOTKELIGTFYGVVAVRNRREIATKRIREGCEDKKTERIYQITKASHVN 65
QY 89 IVKLYGACLNPLVC--LVMEYAGGSLYNVLHGAEPLPYTAHAMSWCLQCSQGVAYLHS 146
Db 66 IVELYGTSRHEGCALLLMEFVDGGSLSFLH-AKSPSYSHAHAFNWAHQIAGGIAYLHG 124
QY 147 MQPKALIHRLDKPPNLLVAGGVTLKIDFGTACDIOHTMTNKGSAWMAPE----- 199
Db 125 MQPKAVIHRLDKPNTLLCEKGLKLIKIDFGTVVDUSQISNAGTCRYKAEVRELDFP 184
QY 200 -----VPEGSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMVAHV 244
Db 185 KSNRIINQPTGFKVLQGNKPKCDVYSWAIITFWELLSRKEPPEQY-NTLIFELYMAIN 243
QY 245 NGTR 248
|

```

```

Db 244 EGKR 247
RESULT 13
AAM25322
ID   AAM25322 standard; Protein; 473 AA.
XX   AC
XX   AAM25322;
XX   DT
XX   16-OCT-2001 (first entry)
XX   DE
XX   Human protein sequence SEQ ID NO:837.
XX   KW
XX   Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
XX   antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
XX   antibacterial; endocrine; cardiant; central nervous system; virucide;
XX   anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
XX   antiaggregant; haemostatic; vulnery; antiulcer; osteopathic; eczema;
XX   dermatological; antiallergic; antisthmatic; antidiabetic; cytostatic;
XX   neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
XX   immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
XX   immanaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
XX   cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
XX   genetic disease; haematopoietic disorder; platelet disorder; asthma;
XX   thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
XX   allergic rhinitis; diabetes; multiple sclerosis; depression;
XX   Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
XX   neurological disorder.
XX   KW
XX   Homo sapiens.
XX   OS
XX   WO200153455-A2.
XX   PN
XX   26-JUL-2001.
XX   PD
XX   22-DEC-2000; 2000WO-US35017.
XX   PF
XX   23-DEC-1999; 99US-0471275.
XX   PR
XX   21-JAN-2000; 2000US-0488725.
XX   PR
XX   25-APR-2000; 2000US-0552317.
XX   PA
XX   (HYSE-) HYSEQ INC.
XX   PI
XX   Tang YT, Liu C, Drmanac RT;
XX   WPI; 2001-457603/49.
XX   DR
XX   N-PSDB; AAM99263.
XX   PT
XX   Isolated human polynucleotides encoding polypeptides, useful for the
XX   treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
XX   PS
XX   Claim 20; Page 191; 1217pp; English.
XX   CC
XX   AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
XX   AAM25963. The proteins can have activities based on the tissues and
XX   cells they are expressed in, such as: antinflammatory; antirheumatic;
XX   antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;
XX   central nervous system; virucide; anti-HIV; fungicide; antimutagen;
XX   cardiovascular; antianaemic; antiaggregant; haemostatic; vulnery;
XX   antiulcer; osteopathic; dermatological; antiallergic; antisthmatic;
XX   antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
XX   antiparkinsonian; and immunostimulant. The proteins and polynucleotides
XX   encoding them can be used in gene therapy, antisense therapy and vaccine
XX   production. The proteins and polynucleotides are useful for screening for
XX   agonists or antagonists of a protein and for the treatment and diagnosis
XX   of disorders associated with the activity of a protein e.g. inflammation,
XX   rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
XX   neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
XX   infections, autoimmunity, genetic diseases, haematopoietic disorders,
XX   anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
XX   osteoporosis, severe combined immunodeficiency, eczema, allergic
XX   rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
XX   Alzheimer's disease, Parkinson's disease, neurodegenerative and

```


CC other developmental disorders of the central nervous system,
CC neuromuscular disorders, myasthenia gravis, cell proliferative disorders
CC such as actinic keratosis, arteriosclerosis, atherosclerosis and cancer
CC including leukaemia, melanoma, myeloma and cancer of the adrenal gland,
CC bladder, bone, bone marrow, liver, lung, muscle, ovary, autoimmune/
CC inflammatory disorder such as Addison's disease, acquired
CC immunodeficiency diseases, allergies, bronchitis, diabetes mellitus,
CC rheumatoid arthritis, microbial infection and trauma.
XX
SQ Sequence 455 AA;

Query Match 16.0%; Score 481; DB 21; Length 455;
Best Local Similarity 30.4%; Pred. No. 3.3e-31;
Matches 132; Conservative 81; Mismatches 151; Indels 70; Gaps 18;
QY 27 NFEEDYKEIEVEVGRGAFGVCKAKW--RAKDVAIK---QIESEERKAFIVELRQL 81
DB 7 SFVQIKFDLQFFENCNGSGSGSVYRAKWISQDKEVAVKLLKIEKAE-----IL 57
QY 82 SRVNHPIVILYACINLP--VCLVMEYAEGLSYLVNHLG--AEPLPYTAAHMSWCILQC 137
DB 58 SVLSHRNIIQFYGVILEPPNYGIVTEYASGLSYLDYINSRSEEM---DMDHMTWATDV 114
QY 138 SQGVAYLHSMQPKALHRLDKPNNLLVAGTVLKIICDFTACDIQTHMTNNK--GSAAW 195
DB 115 AKGMHYLHMEAPVKVTHRLDKSRNVVIAADG-VLKICDPG-ASRFNHTTHMSLVGTTPW 172
QY 196 MAPEVPEGSNYSEKCDVFSWGIILMEVITRRKPFDEIGGPAFRIMW-AVHNGTRPPLIKN 254
DB 173 MAPEVIQSLPVSETCDITYSGVVLMEMLTREVPFKGLEG--LQVAVLVEKNERLTIPSS 230
QY 255 LKPIESLMTRCWSKDPORSGMEEIVKIMTHLMRYFPFGADEPLOYPCQYSDGOSNSAT 314
DB 231 CPRSPFELLHQWEADAKRPFKQIISIL-----ESMSNDT- 267
QY 315 STGSFMDIASNTSNKSDTNMEQVPATNDTIKRLSKLLKNQAKQOSESGRLSL---GAS 371
DB 268 ---SLPDKCNSFLHNKAERCE-IEATLERLKLRLDLSPKEQELKERERRLKNWEOKLT 323
QY 372 HGSSVESLPPTS-----EGKRMADMSIEARIAATT-GNGQPRRRSIQDLTVTGT 421
DB 324 EOSNTPLLLPLAARMSSESYFESKTEESNAEMSQCITATSNGBGHGMNPSLQAWMLMGF 383
QY 422 EPQVSRSSSPSV 435
DB 384 --GDIFSMNKAGAV 395

RESULT 15
AAY83278
ID AAY83278 standard; Protein; 455 AA.
XX
AC AAY83278;
XX
DT 16-AUG-2000 (first entry)
XX
DE Human survival regulating kinase (SRK).
XX
KW Survival regulating kinase; SRK; Raf; MBP; BAD; MAPKK; MAPKK;
KW HAX-1; protein kinase; autophosphorylation; cell growth; regulation;
KW apoptosis; cell survival; nuclear targeting; tumour; human;
KW autoimmune disease.
XX
OS Homo sapiens.
XX
PN WO200022142-A2.
XX
PD 20-APR-2000.
XX
PF 20-SEP-1999; 99WO-US22008.
XX
PR 13-OCT-1998; 98US-0104088.
XX

PA (ONYX-) ONYX PHARM INC.
XX
PI Ruggieri R, Callow M, Diaz P;
XX
XX WPI; 2000-317994/27.
DR N-PSDB; AA293783.
XX
XX Novel human survival regulating kinase polypeptide for screening agents
PT which modulate biological pathways associated with SRK useful in
PT treating autoimmune diseases, tumors and apoptosis-related disorders
XX
XX Claim 4; Figure 2; 62pp; English.
XX
XX Survival regulating kinases (SRK) are a class of proteins involved in
CC cell signal transduction pathways such as mitogen-activated protein
CC kinase pathways. A protein kinase activity means that the SRK can
CC catalyse a reaction in which a phosphate group is transferred from a
CC phosphate donor to a phosphate acceptor amino acid residue,
CC preferably the hydroxyl side chain of a serine or threonine.
CC Substrates for SRK include SRK, MBP and BAD and SRK's protein kinase
CC activity is similar to that of a MAPKK such as Raf. has a range of
CC other activities including a cell growth-regulatory activity, a cell
CC survival promoting activity, a HAX-1 binding activity, an apoptosis
CC suppressing activity a MAPKK activation or stimulatory activity, a
CC nuclear targeting activity and a SRK-specific immunogenic activity.
CC SRK is useful for identifying agents which modulate cellular
CC transformations mediated by Ras and SRK and agents that modulate the
CC apoptosis suppression activity of SRK. This information may be useful
CC in the treatment of autoimmune diseases, tumours and apoptosis
CC related disorders.
XX
SQ Sequence 455 AA;

Query Match 16.0%; Score 481; DB 21; Length 455;
Best Local Similarity 30.4%; Pred. No. 3.3e-31;
Matches 132; Conservative 81; Mismatches 151; Indels 70; Gaps 18;
QY 27 NFEEDYKEIEVEVGRGAFGVCKAKW--RAKDVAIK---QIESEERKAFIVELRQL 81
DB 7 SFVQIKFDLQFFENCNGSGSGSVYRAKWISQDKEVAVKLLKIEKAE-----IL 57
QY 82 SRVNHPIVILYACINLP--VCLVMEYAEGLSYLVNHLG--AEPLPYTAAHMSWCILQC 137
DB 58 SVLSHRNIIQFYGVILEPPNYGIVTEYASGLSYLDYINSRSEEM---DMDHMTWATDV 114
QY 138 SQGVAYLHSMQPKALHRLDKPNNLLVAGTVLKIICDFTACDIQTHMTNNK--GSAAW 195
DB 115 AKGMHYLHMEAPVKVTHRLDKSRNVVIAADG-VLKICDPG-ASRFNHTTHMSLVGTTPW 172
QY 196 MAPEVPEGSNYSEKCDVFSWGIILMEVITRRKPFDEIGGPAFRIMW-AVHNGTRPPLIKN 254
DB 173 MAPEVIQSLPVSETCDITYSGVVLMEMLTREVPFKGLEG--LQVAVLVEKNERLTIPSS 230
QY 255 LKPIESLMTRCWSKDPORSGMEEIVKIMTHLMRYFPFGADEPLOYPCQYSDGOSNSAT 314
DB 231 CPRSPFELLHQWEADAKRPFKQIISIL-----ESMSNDT- 267
QY 315 STGSFMDIASNTSNKSDTNMEQVPATNDTIKRLSKLLKNQAKQOSESGRLSL---GAS 371
DB 268 ---SLPDKCNSFLHNKAERCE-IEATLERLKLRLDLSPKEQELKERERRLKNWEOKLT 323
QY 372 HGSSVESLPPTS-----EGKRMADMSIEARIAATT-GNGQPRRRSIQDLTVTGT 421
DB 324 EOSNTPLLLPLAARMSSESYFESKTEESNAEMSQCITATSNGBGHGMNPSLQAWMLMGF 383
QY 422 EPQVSRSSSPSV 435
DB 384 --GDIFSMNKAGAV 395

Search completed: December 9, 2002, 22:47:16
Job time : 63.9474 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 9, 2002, 22:47:24 ; Search time 33.1468 Seconds
(without alignments)
1679.251 Million cell updates/sec

Title: US-09-830-144-2
Perfect score: 3014
Sequence: 1 MSTASAASSSSSSSAGEMIE.....QCKQLEVRSQQKRGQTS 579

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 73.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	3006	99.7	579	2 JC5955	transforming growt
2	2982.5	99.0	606	2 JC5956	transforming growt
3	2750	91.2	567	2 JC5957	transforming growt
4	493.5	16.4	954	1 S68178	mixed-lineage prot
5	459.5	15.2	847	1 A53800	probable protein k
6	434.5	14.4	407	2 G84635	protein kinase hom
7	434	14.4	412	2 T10671	protein kinase hom
8	433.5	14.4	1338	2 T18287	protein-tyrosine k
9	427.5	14.2	391	2 T48115	protein kinase ATM
10	421.5	14.0	357	2 C84856	probable protein k
11	418.5	13.9	394	2 JU0229	mixed-lineage prot
12	416	13.8	328	2 T16747	hypothetical prote
13	414	13.7	410	2 B35670	protein-tyrosine k
14	412	13.7	668	2 JC2363	protein kinase (EC
15	412	13.7	1030	2 F96763	hypothetical prote
16	411.5	13.7	546	2 D84555	probable protein k
17	411.5	13.7	553	2 T04683	hypothetical prote
18	408.5	13.6	888	2 A55318	serine/threonine p
19	406.5	13.5	370	2 T46150	protein kinase ATN
20	404.5	13.4	888	2 JC5399	dual leucine zippe
21	398	13.2	821	2 T48400	serine/threonine-p
22	397.5	13.2	829	2 T07406	probable protein k
23	391.5	13.0	736	2 T05137	protein kinase hom
24	391	13.0	364	2 G71410	probable protein k
25	391	13.0	387	2 T22511	hypothetical prote
26	390	12.9	963	2 T09911	probable serine/th
27	389.5	12.9	356	2 S61766	protein kinase ATN
28	387	12.8	475	2 T12955	probable protein k
29	385	12.8	438	2 C86273	protein kinases ho

30	383	12.7	982	2 T06576	probable protein k
31	380.5	12.6	1094	2 S49313	protein kinase - s
32	379.5	12.6	1584	2 T18276	protein-tyrosine k
33	376.5	12.5	445	2 T31581	hypothetical prote
34	374.5	12.4	405	2 T52626	probable mitogen-a
35	374	12.4	1015	2 T00726	probable serine/th
36	373	12.4	462	2 S29851	protein kinase 6 (
37	370	12.3	1029	2 H86179	hypothetical prote
38	369.5	12.3	855	2 T20082	hypothetical prote
39	368	12.2	690	2 C96572	protein F12M16.4 l
40	367.5	12.2	1064	1 S57450	protein-tyrosine k
41	366.5	12.2	390	2 T01451	protein kinase hom
42	366	12.1	545	2 T05675	hypothetical prote
43	365.5	12.1	356	2 T48206	protein kinase ATN
44	365.5	12.1	630	1 T01380	protein-tyrosine k
45	365.5	12.1	631	2 I56997	protein-tyrosine k

ALIGNMENTS

RESULT 1

JC5955

transforming growth factor-beta activated kinase (EC 2.7.-.-) 1a - human

C:Species: Homo sapiens (man)

C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000

C:Accession: JC5955

R:Sakurai, H.; Shigemori, N.; Hasegawa, K.; Sugita, T.

Biochem. Biophys. Res. Commun. 243, 545-549, 1998

A>Title: TGF-beta-activated kinase 1 stimulates NF-kappaB activation by an NF-kappaB-ind

A:Reference number: JC5955; MUID:98153801; PMID:9480845

A:Accession: JC5955

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-579 <SAK>

A:Cross-references: DDBJ:AB009356; NID:G2924623; PIDN:BA25025.1; PID:G2924624

C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog

C:Keywords: phosphotransferase

Query Match	99.7%	Score 3006;	DB 2;	Length 579;
Best Local Similarity	99.8%	Pred. No. 1.9e-126;		
Matches 578;	Conservative	0; Mismatches 1;	Indels	0; Gaps
0;				
Qy	1	MSTASAASSSSSSSAGEMIEAPSVLNPFEEIDYKEIEVEEVVGRGAFGVCKAKWRAKDV	60	
Db	1	MSTASAASSSSSSSAGEMIEAPSVLNPFEEIDYKEIEVEEVVGRGAFGVCKAKWRAKDV	60	
Qy	61	AIKQIESESEKAFIVELRQLSRVNHNPVVKLYGACLNVPCLVMVEYAGGSLYNVLHGAE	120	
Db	61	AIKQIESESEKAFIVELRQLSRVNHNPVVKLYGACLNVPCLVMVEYAGGSLYNVLHGAE	120	
Qy	121	PLPYVYTAHAHMSWCLQCSQGVAYLHSMOPKALIHRLDKPPNLLLVAGTVLKI CDGFTAC	180	
Db	121	PLPYVYTAHAHMSWCLQCSQGVAYLHSMOPKALIHRLDKPPNLLLVAGTVLKI CDGFTAC	180	
Qy	181	DIQTHMTNKNKSAAMWAPFEGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIM	240	
Db	181	DIQTHMTNKNKSAAMWAPFEGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIM	240	
Qy	241	WAVHNGTPPLIKNLPKPIESIMTRCWSKQSPSMEIEVKIMTHLMRYPPGGADEPLQY	300	
Db	241	WAVHNGTPPLIKNLPKPIESIMTRCWSKQSPSMEIEVKIMTHLMRYPPGGADEPLQY	300	
Qy	301	PCQYSDGQSNATSGSFMDIATNTSNKSDTNMQVPATNDTIKRLSKLLKNQAKQ	360	
Db	301	PCQYSDGQSNATSGSFMDIATNTSNKSDTNMQVPATNDTIKRLSKLLKNQAKQ	360	
Qy	361	SESGRLSICAGSHGSSVESLPPTSGKMSADMSIEARIAATTGNGOPRRRSIODLTVTG	420	
Db	361	SESGRLSICAGSHGSSVESLPPTSGKMSADMSIEARIAATTGNGOPRRRSIODLTVTG	420	
Qy	421	TEPGQVSRSSSPSVRMITTSQPTSEKPTRSHPTDSDTNGSDNSIPMAYITLDHQL	480	
Db	421	TEPGQVSRSSSPSVRMITTSQPTSEKPTRSHPTDSDTNGSDNSIPMAYITLDHQL	480	

Db 421 TEEGQVSSRSSSVMTTSGPTSEKPTRSHWTPDDSTDTNGSNSIPMAVLTLDHQL 480
QY 481 QPLAPCNSKESNAVEHQCKMAQOEYMKVQTEITALLLQKQELVAELDDEKQNTSRL 540
Db 481 QPLAPCNSKESNAVEHQCKMAQOEYMKVQTEITALLLQKQELVAELDDEKQNTSRL 540
QY 541 VOEHKKLLDENKSLSTYYQCKQKQLEVIRSOQCKRGTS 579
Db 541 VOEHKKLLDENKSLSTYYQCKQKQLEVIRSOQCKRGTS 579

RESULT 2

JC5956
transforming growth factor-beta activated kinase (EC 2.7.-.-) 1b - human
C/Species: Homo sapiens (man)
C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C/Accession: JC5956
R/Sakurai, H.; Shigemori, N.; Hasegawa, K.; Sugita, T.
Biochem. Biophys. Res. Commun. 243, 545-549, 1998
A/Title: TGF-beta-activated kinase 1 stimulates NF-kappaB activation by an NF-kappaB-ind
A/Reference number: JC5955; MUID:98153801; PMID:9480845
A/Accession: JC5956
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-606 <SAK>
A/Cross-references: DDBJ:AB009357; NID:92924625; PID:BA25026.1; PID:92924626
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C/Keywords: phosphotransferase

Query Match 99.0%; Score 2982.5; DB 2; Length 606;
Best Local Similarity 95.4%; Pred. No. 2.2e-125;
Matches 578; Conservative 0; Mismatches 1; Indels 27; Gaps 1;
QY 1 MSTASASSSSSSAGEMIEAPSQVLNFEEDYKEIEVEEVVGRGAFGVCAKMRKDV 60
Db 1 MSTASASSSSSSAGEMIEAPSQVLNFEEDYKEIEVEEVVGRGAFGVCAKMRKDV 60
QY 61 AIKQIESSEKKAIVELRQLSRVNHNPNIYKLYGACLNPCVLMVEYAEAGSLYVNLHGA 120
Db 61 AIKQIESSEKKAIVELRQLSRVNHNPNIYKLYGACLNPCVLMVEYAEAGSLYVNLHGA 120
QY 121 PLPYTTAAHMSWCLQSGQVAYLHSMQPAKLIHRDLKPNLLLVAGGVTLKICDEGTAC 180
Db 121 PLPYTTAAHMSWCLQSGQVAYLHSMQPAKLIHRDLKPNLLLVAGGVTLKICDEGTAC 180
QY 121 PLPYTTAAHMSWCLQSGQVAYLHSMQPAKLIHRDLKPNLLLVAGGVTLKICDEGTAC 180
Db 121 PLPYTTAAHMSWCLQSGQVAYLHSMQPAKLIHRDLKPNLLLVAGGVTLKICDEGTAC 180
QY 181 DIQHTMTNKKSAAMMAPEVEFGSNSEKCDVFSWGIIIMEVITRRKPFDEIGGPAFRIM 240
Db 181 DIQHTMTNKKSAAMMAPEVEFGSNSEKCDVFSWGIIIMEVITRRKPFDEIGGPAFRIM 240
QY 241 WAVHNGTRPPLIKNLPKPIESLMTRCWSKQPSQPSMEIYKIMTHLMRYFPGADEPLQY 300
Db 241 WAVHNGTRPPLIKNLPKPIESLMTRCWSKQPSQPSMEIYKIMTHLMRYFPGADEPLQY 300
QY 301 PCQYSDGQNSATSTGSPMDIASTNTSNKSDTNMEQVPATNDTIKRLSKLLKNOAKQ 360
Db 301 PCQYSDGQNSATSTGSPMDIASTNTSNKSDTNMEQVPATNDTIKRLSKLLKNOAKQ 360
QY 301 PCQYSDGQNSATSTGSPMDIASTNTSNKSDTNMEQVPATNDTIKRLSKLLKNOAKQ 360
Db 301 PCQYSDGQNSATSTGSPMDIASTNTSNKSDTNMEQVPATNDTIKRLSKLLKNOAKQ 360
QY 361 SESGRSLIGASHGSSVESLPTSEGRKMSADMSEIEARIATYAKPKRGHKRTASFGN 420
Db 361 SESGRSLIGASHGSSVESLPTSEGRKMSADMSEIEARIATYAKPKRGHKRTASFGN 420
QY 404 -----GNQPPRRRSIODLTVTGTEPGQVSSRSSSPSVMTTSGPTSEKPTRSH 453
Db 404 -----GNQPPRRRSIODLTVTGTEPGQVSSRSSSPSVMTTSGPTSEKPTRSH 453
QY 421 ILDPVEIVISGNQPPRRRSIODLTVTGTEPGQVSSRSSSPSVMTTSGPTSEKPTRSH 480
Db 421 ILDPVEIVISGNQPPRRRSIODLTVTGTEPGQVSSRSSSPSVMTTSGPTSEKPTRSH 480
QY 454 WTPDDSTDTNGSDNSIPMAVLTLDHQLQPLAPCNSKESNAVEHQCKMAQOEYMKVQTEI 513
Db 454 WTPDDSTDTNGSDNSIPMAVLTLDHQLQPLAPCNSKESNAVEHQCKMAQOEYMKVQTEI 513
QY 481 WTPDDSTDTNGSDNSIPMAVLTLDHQLQPLAPCNSKESNAVEHQCKMAQOEYMKVQTEI 540
Db 481 WTPDDSTDTNGSDNSIPMAVLTLDHQLQPLAPCNSKESNAVEHQCKMAQOEYMKVQTEI 540
QY 514 ALLQKQELVAELDDEKQNTSRLVOEHKKLLDENKSLSTYYQCKQKQLEVIRSOQ 573
Db 514 ALLQKQELVAELDDEKQNTSRLVOEHKKLLDENKSLSTYYQCKQKQLEVIRSOQ 573
QY 541 ALLQKQELVAELDDEKQNTSRLVOEHKKLLDENKSLSTYYQCKQKQLEVIRSOQ 600
Db 541 ALLQKQELVAELDDEKQNTSRLVOEHKKLLDENKSLSTYYQCKQKQLEVIRSOQ 600

QY 574 KRQGTG 579
Db 601 KRQGTG 606

RESULT 3

JC5957
transforming growth factor-beta activated kinase (EC 2.7.-.-) 1c - human
C/Species: Homo sapiens (man)
C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 17-Mar-2000
C/Accession: JC5957
R/Sakurai, H.; Shigemori, N.; Hasegawa, K.; Sugita, T.
Biochem. Biophys. Res. Commun. 243, 545-549, 1998
A/Title: TGF-beta-activated kinase 1 stimulates NF-kappaB activation by an NF-kappaB-ind
A/Reference number: JC5955; MUID:98153801; PMID:9480845
A/Accession: JC5957
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-567 <SAK>
A/Cross-references: DDBJ:AB009358
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C/Keywords: phosphotransferase

Query Match 91.2%; Score 2750; DB 2; Length 567;
Best Local Similarity 88.9%; Pred. No. 4e-115;
Matches 539; Conservative 0; Mismatches 1; Indels 66; Gaps 2;

QY 1 MSTASASSSSSSAGEMIEAPSQVLNFEEDYKEIEVEEVVGRGAFGVCAKMRKDV 60
Db 1 MSTASASSSSSSAGEMIEAPSQVLNFEEDYKEIEVEEVVGRGAFGVCAKMRKDV 60
QY 61 AIKQIESSEKKAIVELRQLSRVNHNPNIYKLYGACLNPCVLMVEYAEAGSLYVNLHGA 120
Db 61 AIKQIESSEKKAIVELRQLSRVNHNPNIYKLYGACLNPCVLMVEYAEAGSLYVNLHGA 120
QY 121 PLPYTTAAHMSWCLQSGQVAYLHSMQPAKLIHRDLKPNLLLVAGGVTLKICDEGTAC 180
Db 121 PLPYTTAAHMSWCLQSGQVAYLHSMQPAKLIHRDLKPNLLLVAGGVTLKICDEGTAC 180
QY 121 PLPYTTAAHMSWCLQSGQVAYLHSMQPAKLIHRDLKPNLLLVAGGVTLKICDEGTAC 180
Db 121 PLPYTTAAHMSWCLQSGQVAYLHSMQPAKLIHRDLKPNLLLVAGGVTLKICDEGTAC 180
QY 181 DIQHTMTNKKSAAMMAPEVEFGSNSEKCDVFSWGIIIMEVITRRKPFDEIGGPAFRIM 240
Db 181 DIQHTMTNKKSAAMMAPEVEFGSNSEKCDVFSWGIIIMEVITRRKPFDEIGGPAFRIM 240
QY 241 WAVHNGTRPPLIKNLPKPIESLMTRCWSKQPSQPSMEIYKIMTHLMRYFPGADEPLQY 300
Db 241 WAVHNGTRPPLIKNLPKPIESLMTRCWSKQPSQPSMEIYKIMTHLMRYFPGADEPLQY 300
QY 301 PCQYSDGQNSATSTGSPMDIASTNTSNKSDTNMEQVPATNDTIKRLSKLLKNOAKQ 360
Db 301 PCQYSDGQNSATSTGSPMDIASTNTSNKSDTNMEQVPATNDTIKRLSKLLKNOAKQ 360
QY 301 PCQYSDGQNSATSTGSPMDIASTNTSNKSDTNMEQVPATNDTIKRLSKLLKNOAKQ 360
Db 301 PCQYSDGQNSATSTGSPMDIASTNTSNKSDTNMEQVPATNDTIKRLSKLLKNOAKQ 360
QY 361 SESGRSLIGASHGSSVESLPTSEGRKMSADMSEIEARIATYAKPKRGHKRTASFGN 420
Db 361 SESGRSLIGASHGSSVESLPTSEGRKMSADMSEIEARIATYAKPKRGHKRTASFGN 420
QY 404 -----GNQPPRRRSIODLTVTGTEPGQVSSRSSSPSVMTTSGPTSEKPTRSH 453
Db 404 -----GNQPPRRRSIODLTVTGTEPGQVSSRSSSPSVMTTSGPTSEKPTRSH 453
QY 421 ILDPVEIVISGNQPPRRRSIODLTVTGTEPGQVSSRSSSPSVMTTSGPTSEKPTRSH 480
Db 421 ILDPVEIVISGNQPPRRRSIODLTVTGTEPGQVSSRSSSPSVMTTSGPTSEKPTRSH 480
QY 454 WTPDDSTDTNGSDNSIPMAVLTLDHQLQPLAPCNSKESNAVEHQCKMAQOEYMKVQTEI 513
Db 454 WTPDDSTDTNGSDNSIPMAVLTLDHQLQPLAPCNSKESNAVEHQCKMAQOEYMKVQTEI 513
QY 514 ALLQKQELVAELDDEKQNTSRLVOEHKKLLDENKSLSTYYQCKQKQLEVIRSOQ 573
Db 514 ALLQKQELVAELDDEKQNTSRLVOEHKKLLDENKSLSTYYQCKQKQLEVIRSOQ 573
QY 509 -----QELVAELDDEKQNTSRLVOEHKKLLDENKSLSTYYQCKQKQLEVIRSOQ 561
Db 509 -----QELVAELDDEKQNTSRLVOEHKKLLDENKSLSTYYQCKQKQLEVIRSOQ 561
QY 574 KRQGTG 579
Db 562 KRQGTG 567

RESULT 4

568178 mixed-lineage protein kinase 2 (EC 2.7.1.1-) - human
C:Species: Homo sapiens (man)
C:Date: 28-Oct-1996 #sequence revision 13-Mar-1997 #text_change 11-Jun-1999
C:Accession: S68178; I38044; S32468
R:Dorow, D.S.; Devereux, L.; Tu, G.F.; Price, G.; Nicholl, J.K.; Sutherland, G.R.; Simpson, J.R.; Biochem. 234, 492-500, 1995
E:J. Biochem. 234, 492-500, 1995
A:Title: Complete nucleotide sequence, expression, and chromosomal localisation of human mixed-lineage protein kinase 2
A:Reference number: S68178; MUID:96128179; PMID:8536694
A:Accession: S68178
A:Molecule type: mRNA
A:Residues: 1-954 <DOR>
A:Cross-references: EMBL:X90846; NID:g971419; PIDN:CAA62351.1; PID:g971420
R:Kaoh, M.; Hirai, M.; Sugimura, T.; Terada, M.
Oncogene 10, 1447-1451, 1995
A:Title: Cloning and characterization of MST, a novel (putative) serine/threonine kinase
A:Reference number: I38044; MUID:95249256; PMID:7731697
A:Accession: I38044
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-461,'A','V', 465-470,'S',472-806,'R', 808-817,'A',819-954 <RES>
A:Cross-references: EMBL:Z48615; NID:g758592; PIDN:CAA88531.1; PID:g758593
R:Dorow, D.S.; Devereux, L.; Dietzsch, E.; de Kretser, T.
Eur. J. Biochem. 213, 701-710, 1993
A:Title: Identification of a new family of human epithelial protein kinases containing a SH3 domain
A:Reference number: S32467; MUID:93238756; PMID:8477742
A:Accession: S32468
A:Molecule type: mRNA
A:Residues: 244-464,'AQAAGRQHPQALWL' <DQ2>
C:Genetics:
A:Gene: GDB:MLK2; GDB:MST
A:Cross-references: GDB:362654; GDB:624810; OMIM:600137
A:Map position: 19q13.1-19q13.2
C:Superfamily: mixed-lineage protein kinase 2; protein kinase homology; SH3 homology
C:Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific protein kinase
F:23-76/Domain: SH3 homology <SH3>
F:96-364/Domain: protein kinase homology <KIN>
F:104-112/Region: protein kinase ATP-binding motif
F:384-405/Region: leucine zipper motif
F:419-440/Region: leucine zipper motif
F:449-463/Region: basic
F:125,145,222,224/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 16.4%; Score 493.5; DB 1; Length 954;
Best Local Similarity 28.9%; Pred. No. 6e-15;
Matches 145; Conservative 82; Mismatches 192; Indels 83; Gaps 16;

Qy 21 APSQVLPFEEDYKIEVEEVVVGCAFVVCCKAKRAKDVAIKQIESERKAFIV---- 76
Db 83 APAGLQFQEIIPFHELOQEETIGVGGFKVYRALMRGEVAVKAARLDPKDPVATAEQV 142
Qy 77 --ELRLSRVNHNPVTKVLYGACLPN--VCLWVEYAGGSLYNVLHGAEPYPYTAHAAMS 132
Db 143 CQEARLFCALGHPNIIALRGACLPNPHCLWVEYAGGALSRLVLAGRVPVPHV----LVN 198
Qy 133 WLCQSQGVAYLHSMQPKALIHRLKPNLLV-----AGGTVLKICDFGTACD-IQT 184
Db 199 WAVQVARGMNYLHNDAPVPIIHRLDKSINILILEAIEHNHLADTVLKITDFGLAREWHKT 258
Qy 185 HMTNKGSAAMWAEVFGSGNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMNAV- 243
Db 259 TKMSAAGTYAWMAPEVIRLSLFSKSSDWSFGVLLWELLTGEVPYREI--DALAVAYGVA 316
Qy 244 HNGTRPPLIKNLPKPIESLMTRCRCKSPQSPSMEETVKIMTHL-----MRYFPGA 294
Db 317 MNKLTLPSTPCPEFPARLLCECDPDHPGDPDFGSLIKRLEVIEQSQALFOMPLESFSL 376
Qy 295 DEPLQYPCQY-----SDEGOSNATSTGSP-----MDIAT 325
Db 377 QEDWKLEIQHMFDDLRTKEKELRSREBELLRAAEQRFQEBQRLRRRQELAEKREWDI--- 433
Qy 326 NTSNKSNTMNEQVATNDTIKRLSKLLKNQAKQOQSSG----RUSLGASHGSSVESLPPT 382

Matches 98; Conservative 47; Mismatches 100; Indels 23; Gaps 8;

OY 35 EIEVEEVGRGAPGVVCAKMPAKDAVK--QIESSEKAFIVELRQ---LSRVNHPN 88
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 2 ETLLEEIIIGGFQKVYRAFWIGDEVAAYKARHDPDEDISQTIEVRQFAKLKHPN 61

OY 89 IYKLVGACLNLP--VCLVMWYAAGGSLYNVLHGAEPLPYTAHANSMWLCSOGQAVYHS 146
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 62 IIALRCVLCKEPNLCLWMFARAGPILNRVLVSGKRIPDI---LVNMAQLARGNNYLHD 117

OY 147 MOPKALIHRLDKPPLLNV-----AGGTVLKICDFGTACD-IQTHMTNNKGSAAAMAP 198
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 118 KXIVPIIHRDLKSSNLILLQKVENGLSNKLIKIDFGLAREMHNHTTKMSAAGTANMAP 177

OY 199 EVFEEGNSSEKCDVFSWGIIIMWEVITRRKPPEDIGPAFRIMAV-HNGTRPPLIKNLPK 257
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 178 EVIRASMFESKSDVMSYGVLMLLELLTGVEVFRGIDG--LRVAVGANMKLALPISTCPE 235

OY 258 PLESIMTRCMSKDPSORSPSMEIIVKIMT 285
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 236 PPAKLMECDMNPDRSPSRFTNILDQLT 263

RESULT 12

T16747
hypothetical protein R1Jf6.7 - Caenorhabditis elegans
CSpecies: Caenorhabditis elegans
CDate: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Mar-2000
CAccession: T16747
R.Miller N.
submitted to the EMBL Data Library April 1994
A:Description: The sequence of C. elegans cosmid R1Jf6.
A:Reference number: Z18570
A:Accession: T16747
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-328 <ML>
A:Cross-references: EMBL:U00046; NID:g470358; PID:g470364; PIDN:AAC47047.1; GSFDB:GN0002
A:Experimental source: strain Bristol NZ; clone R1Jf6
C:Genetics:
A:Gene: CESP.R1Jf6.7
A:Map position: 3
A:Introns: 20/3; 160/3; 222/2; 286/2
C:Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3 h

Query Match 13.8%; Score 416; DB 2; Length 328;
Best Local Similarity 33.4%; Pred. No.5.5e-12;
Matches 99; Conservative 57; Mismatches 112; Indels 28; Gaps 11;

OY 1 MSTASAASSSSSSAGEMIAPSOVLNBEIDYKEIEVEEVGRGAPGVVCAKMPAKDV 60
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 22 MSTPTNSESTSSSSNNNS----DQRVLFPDIORDDIOVDGHIGVTFGAVFSGNWTLPDG 76

OY 61 AIKOIESEBERKAFLVELRQ--LSRVNHNTIYKLYGACL--NPLCLMEVAEAGSLTNV 115
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 77 SORTI--ALKVFLEKEAEITLSKIRKHNIIFQGIKCANGDNDFITEVAEKESLYDF 133

OY 116 LHGAEPPLPYTTAAHAMS-----WCLCSOGVAYLHSMOPKALIHRLDKPPLLVAAGT 169
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 134 IHSESSQSASSSGSNSPVVVKMASQIASGIQYHYDAVDTIIHRDLKSKNVL-DKNL 192

OY 170 VLKICDFGTACDIQTHMTNK--GSAAMMADE-VFEGSNYSEKCDVFSWGIIIMEVITR 225
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 193 VKCIICDFGSKDL-THCSTAPSWGCTAAMSPSEMILQSEGTLTATDVMSYGVLMELISK 251

OY 226 RKPFPEIGPARIMNAV-HNGTRPPLIKNLPKPLESIMTRCMSKDPSORSPMEI 280
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 252 EVFYKDYK--EFRIFTMITOSGITLAIPPSCCAPLKOLMSNCWKATPPDXRAMROI 305

RESULT 13

protein-tyrosine kinase (EC 2.7.1.112) 2 - sjime mold (Dictyostelium discoideum) (fragme

C:Species: Dictyostelium discoideum
C:Date: 28-Sep-1990 #sequence_revision 28-Sep-1990 #text_change 04-Feb-2000
C:Accession: B35670
R:Tan, J.L.; Spudich, J.A.
Mol. Cell. Biol. 10, 3578-3583, 1990
A:Title: Developmentally regulated protein-tyrosine kinase genes in Dictyostelium discoideum
A:Reference number: A35670; MUID:90287147; PMID:1972546
A:Accession: B35670
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-410 <TAN>
A:Cross-references: GB:K3784; NID:9167777; PIDN:AA33203.1; PID:9167778
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homology <KIN>
C:Keywords: ATP; phosphotransferase; tyrosine-specific protein kinase
F:106-371/Domain: protein kinase homology <KIN>
F:114-122/Region: protein kinase ATP-binding motif

Query Match 13.7%; Score 414; DB 2; Length 410;
Best Local Similarity 31.5%; Pred. No. 8.4e-12;
Matches 104; Conservative 62; Mismatches 124; Indels 40; Gaps 12;

QY 28 FEE-----IDYKEIEVEEVGRGAFGVVCKAKMPAKDVAIKOI---ESESERK 72
DB 89 FEELKSLIGREYIITDINDIOFIQKVEGAFSEWEGWKGIVHAIKLIKIGDEBFKE 148
QY 73 AFIVELRQLSVHNPNVIKLYGACLNPFCLWMEYAEGSLYNVLH---GAEPFLPYTTAAH 129
DB 149 RFIREVQNLKKGNHNIYMFIGACYKPACTIETEMAGSGLYNIIHNPNSTPKVKYSFPL 208
QY 130 AMSNQLQSSQVAVLHSMQPKALIHRLDKPNNLLVAGTVLKICDGTACD--IQTHMT 187
DB 209 VLKANTDALGLLHLHSI---TVHRLDITSQNILLDELGNI-KISDPGLSAEKSRRESGMT 264
QY 188 NNKG---SAAMAAEVEEG-SNYSEKCDVFSWGIILWEVITRRKPFDEIGSPAFRIWAV 243
DB 265 MTNGICNPKRRRPELIRKNGHYSEKVDYICFSLVWEILITGELPSPDLDG-SQRSKQVA 323
QY 244 HNGTRPPLIKNLKPIESLMTRCWSKDPQSQRPSNEIYIKMTILMRYPGADPLQYPCQ 303
DB 324 YAGIRPRPIPECYDELKILLTQCWEADPNDRPPTYIVNKLKEI-----SMNPDIGP--- 375
QY 304 YSDEGQSNATSITGSPMDIASTNTSNKSDT 333
DB 376 VSDQFYQYSEPTPR---LALSNQSSNSS 402

RESULT 14
JC2363
protein kinase (EC 2.7.1.37) ZPK - human
N:Alternate names: leucine-zipper protein kinase
C:Species: Homo sapiens (man)
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 02-Feb-2001
C:Accession: JC2363
R:Reddy, U.R.; Pleasure, D.
Biochem. Biophys. Res. Commun. 202, 613-620, 1994
A:Title: Cloning of a novel putative protein kinase having a leucine zipper domain from
A:Reference number: JC2363; MUID:94311945; PMID:8037767
A:Accession: JC2363
A:Molecule type: mRNA
A:Residues: 1-668 <RED>
A:Cross-references: EMBL:U07358
A:Experimental source: brain
A:Note: the nucleotide sequence for this amino acid sequence is inconsistent with that
he codon ACC for residue 661 as Pro, the codon GAAACACCTCTTCA for residues 664-668 as
C:Comment: This protein belongs to the family of non-receptor kinase.
C:Genetics:
A:Gene: GDB:ZPK
A:Cross-references: GDB:383963; OMIM:600447
A:Map position: 12q13-12q13
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homology
C:Keywords: ATP; leucine zipper; nucleotide binding; P-loop; phosphotransferase
F:123-371/Domain: protein kinase ATP-binding motif
F:131-139/Region: protein kinase ATP-binding motif

F;443-471/Region: leucine zipper motif
F;538-545/Region: nucleotide-binding motif A (p-loop)
F;152/Active site: Lys #status predicted

Query Match 13.7%; Score 412; DB 2; Length 668;
Best Local Similarity 24.6%; Pred. No. 1.7e-11;
Matches 133; Conservative 79; Mismatches 181; Indels 148; Gaps 20;

Qy	30	EIDYKETEVEVGRGAFVGVCKAKWAKOAVIAHQIIESESEKAFIVELRQLQRVNHPI	89
Db	119	EVPPFETLIDQWGSQAQAVFLGRFFGEVAVKVRDLKE-----TDIRHLRKLHPH	173
Qy	90	VKLGYACLNP--VCLVMEYAEAGGLYNVLHGAELPYVYTAAHAMSWCLQCSOGVAYLHSM	147
Db	174	ITFGVCTQAPCYCILMEFCAQGLQYEVLRAGRPV---TPSLLVDWSMGTAGMNYLHLH	230
Qy	148	QPKALIIHRDLKPNLLVAGGTVLKIDFGTACDIOQTHMTNNK--GSAANMAPEVFECSN	205
Db	231	K--IIHRDLKSPN-MLIITYDDVVKISDFGTSKELSDKSTKMSFAGTVANMAPEVIRNEP	286
Qy	206	YSEKCDVFSGIIILWEVITRRKPFDETGGAFAFRIMWAV-HNGTRPPPLIKNLPDIESLMT	264
Db	287	VSEKVIDWSFGVJWELLTGEBIKYDWDSSA--IIMGVGSNLSHLVPSSCPDGFKILLR	344
Qy	265	RCWSKDPQRSMEIEVKIMTHLMRYPPGADEPLQYPCQYSDGCSNSATSTGSEFMDIAS	324
Db	345	QCMWSKPNRPSFRQ---ILUHL-----DIAS	368
Qy	325	TNT-SNKSDDT-----NMEQVPATNDTIKRLBSKLL-----	353
Db	369	ADVLSTQETFKSQAEWRBEVKLHFKEIKSEGCTLHRLBEELVMRRREELRHALDIREH	428
Qy	354	-----KNQAKQOSESGR-----SLGASHGSSVES	378
Db	429	YERKLERANNLYMBELNALMLQLEKERELLREQUALERRCPGLLKPHPSRGLLHGNTMEK	488
Qy	379	L-----PPTSEGKRMADMSIEIARI-----AATTGNGQPRRSIQDLTVTGTBPGQVSS	428
Db	489	LIKRNVPQNLSPHSQRPDIILKASLSLPKLDAULSGVGLP-----GCPKAPPSFGR--\$	540
Qy	429	RSSSPSRMITTSPTSEKP---TRSHPWTPDDSTDYNG-----SDNSINIPMAYLTLDHQ	479
Db	541	RRGKTRHRKASAKSCGDLPLGLRTAVPHEPGPGSPGLGCGGGSAMEACPPALRGLLHD	600
Qy	480	L 480	
Db	601	L 601	

RESULT 15

K85001.15
 F96763
 hypothetical protein F25p22.8 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: F96763
 R:Theologus, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: F96763
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1030 >STO>
 A:Cross-references: GB:AE005173; NID:g692730; PIDN:AAF24836.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: F25p22.8

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model
Run on: December 9, 2002, 22:45:29 ; Search time 62.5512 Seconds
(without alignments)
1907.257 Million cell updates/sec

Title: US-09-830-144-2
Perfect score: 3014
Sequence: 1 MSTASAGSSSSSAGEMIE.....QCKKQLEVRISQQKROGTS 579

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL_21.*
- 1: sp_archaea.*
 - 2: sp_bacteria.*
 - 3: sp_fungi.*
 - 4: sp_human.*
 - 5: sp_invertebrate.*
 - 6: sp_mammal.*
 - 7: sp_mhc.*
 - 8: sp_organelle.*
 - 9: sp_phase.*
 - 10: sp_plant.*
 - 11: sp_rodent.*
 - 12: sp_virus.*
 - 13: sp_vertebrate.*
 - 14: sp_unclassified.*
 - 15: sp_rvirus.*
 - 16: sp_bacteriap.*
 - 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2963.5	98.3	606	11 Q923A8	Q923A8 mus musculus
2	2819	93.5	539	4 Q9NTR1	Q9NTR1 homo sapien
3	2795.5	92.8	566	4 Q9NTR2	Q9NTR2 homo sapien
4	2638.5	87.5	616	13 O73613	O73613 xenopus lae
5	2512	83.3	491	4 Q9NZ70	Q9NZ70 homo sapien
6	2301.5	76.4	478	4 Q9NTR4	Q9NTR4 homo sapien
7	1310	43.5	260	4 Q9UG54	Q9UG54 homo sapien
8	966	32.1	678	5 Q9VJQ6	Q9VJQ6 drosophila
9	505	16.8	252	5 Q9VCV0	Q9VCV0 drosophila
10	481	16.0	455	4 Q9HCC4	Q9HCC4 homo sapien
11	478	15.9	454	11 Q9ESL3	Q9ESL3 mus musculus
12	462.5	15.3	1066	4 Q9H2N5	Q9H2N5 homo sapien
13	461.5	15.2	850	11 Q9JJI5	Q9JJI5 mus musculus
14	459.5	15.2	847	4 Q16584	Q16584 homo sapien
15	457	15.2	1001	11 Q8VDG6	Q8VDG6 mus musculus
16	456.5	15.1	800	4 Q9HDD2	Q9HDD2 homo sapien

Q9hcc5 homo sapien
Q9nyl2 homo sapien
Q9nyl2 homo sapien
Q8wnn1 homo sapien
Q9esl4 mus musculus
Q9xc6 caenorhabdi
Q90zy8 brachydanio
Q9hly7 homo sapien
Q8wnn2 homo sapien
Q9w313 drosophila
Q9svf6 drosophila
Q95un8 drosophila
Q9zq31 arabidopsis
Q9m085 arabidopsis
Q23846 dictyosteli
Q23927 dictyosteli
Q22100 arabidopsis
Q9vw24 drosophila
Q9sim8 arabidopsis
Q9bi25 dictyosteli
Q94c42 triticum ae
Q9y2v6 homo sapien
Q21982 caenorhabdi
Q8ry96 arabidopsis
Q8rw16 arabidopsis
Q9c9u5 arabidopsis
Q22558 arabidopsis
Q81808 arabidopsis
Q8wy25 homo sapien

ALIGNMENTS

RESULT 1

Q923A8 ID Q923A8 PRELIMINARY; PRT; 606 AA.
AC Q923A8; 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Unknown (protein for MGC:5989).
GN MAP3K7.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC006665; AA06665.1; -.
DR MGD; MGI:1348877; Map3k7.
DR InterPro; IPR000719; Euk_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW ATP-binding; Transferase.
SQ SEQUENCE 606 AA; 67194 MW; AB8664F389272102 CRC64;

Query Match 98.3%; Score 2963.5; DB 11; Length 606;
Best Local Similarity 94.7%; Pred. No. 2.4e-193;
Matches 574; Conservative 1; Mismatches 4; Indels 27; Gaps 1;
QY 1 MSTASAGSSSSSAGEMIEAPSVLNFEEDYKEIEVEEVVGRGAFGVCAKRAKDV 60
Db 1 MSTASAGSSSSSSSASEMEAPSVLNFEEDYKEIEVEEVVGRGAFGVCAKRAKDV 60
QY 61 ATKQESSERKAFIVEIQRVRNHPNIVKLYGACLPVCLVMEYAGGSLYNVLHGA 120

DB 61 AIKQIESESRKAFIVELRQLSRVNHNIYKLYGACLNIPVCLVMEYAEGLSYNVLHGA 120
 QY 121 PLPYRTAAHMSWCLQSGVAVLHMQPALLIHRDLKPNLLLVAGTVLKI CDEGTAC 180
 DB 121 PLPYRTAAHMSWCLQSGVAVLHMQPALLIHRDLKPNLLLVAGTVLKI CDEGTAC 180
 QY 181 DICTHMTNNKGSAAAMAPEVFEESNSSEKCDVFSWGIILMEVITRRKPFDEIGGPAFRIM 240
 DB 181 DICTHMTNNKGSAAAMAPEVFEESNSSEKCDVFSWGIILMEVITRRKPFDEIGGPAFRIM 240
 QY 241 MAVHNGTRPPLIKNLPKPIESLMTRCMSKDPQSPSMEEIVIKMTLHMYFPGADEPLQY 300
 DB 241 MAVHNGTRPPLIKNLPKPIESLMTRCMSKDPQSPSMEEIVIKMTLHMYFPGADEPLQY 300
 QY 301 PCQYSDGQNSATSTGSEFMDIASTNTSNKSDTNMEQVPATNDTIKRLSKLLKNQAKQ 360
 DB 301 PCQYSDGQNSATSTGSEFMDIASTNTSNKSDTNMEQVPATNDTIKRLSKLLKNQAKQ 360
 QY 361 SESGRSLGASHGSSVESLPPTSEGRMSADMEIEARLAAT----- 402
 DB 361 SESGRSLGASHGSSVESLPPTSEGRMSADMEIEARLAAT----- 402
 QY 403 -----TGNGQPRRRSIODLTVTGTEPGQVSSRSSSPSVRMITTSSEKPTRSH 453
 DB 421 ILDPPELVISGNQPRRRSIODLTVTGTEPGQVSSRSSSPSVRMITTSSEKPTRSH 480
 QY 454 WTPDDSTDTNGSDNSIPMAVYLLDHOLOPLAPCNSKESMAVFEQCKMAQSEYMKVQTEI 513
 DB 481 WTPDDSTDTNGSDNSIPMAVYLLDHOLOPLAPCNSKESMAVFEQCKMAQSEYMKVQTEI 540
 QY 514 ALLQKQELVAVELDDQEQQNTSRLVQEHKKLLDENKSLSTYYQCKKQLEVIRSQQ 573
 DB 541 ALLQKQELVAVELDDQEQQNTSRLVQEHKKLLDENKSLSTYYQCKKQLEVIRSQQ 600
 QY 574 KRGQTS 579
 DB 601 KRGQTS 606

RESULT 2
 Q9NTR1 PRELIMINARY; PRT; 539 AA.
 ID Q9NTR1;
 AC Q9NTR1;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE D154G14.1.1 (Mitogen-activated protein kinase kinase 7
 (TGF-beta activated kinase 1a (TAK1))) (Fragment).
 GN MAP3K7.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tracey A.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 CC -i- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AL121964; CAB87607.1; -.
 DR HSSP; P12931; 1FMK.
 DR InterPro; IPR000719; Euk_pkinase.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR InterPro; IPR004040; STY_pkinase.
 DR InterPro; IPR01245; Tyr_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Euk_pkinase; 1.
 DR SMART; SM00221; STYKc_1
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00109; TYRKINASE.
 DR ProDom; PD000001; Euk_pkinase; 1.
 DR SMART; SM00221; STYKc_1
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 FT NON_TER 1

SEQ SEQUENCE 539 AA; 6006 MW; E6183F553CC7F324 CRC64;
 Query Match 93.5%; Score 2819; DB 4; Length 539;
 Best Local Similarity 99.8%; Pred. No. 1,4e-183;
 Matches 538; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 41 VVGKAGFVVCWKAKRAKDVAIKQIESESRKAFIVELRQLSRVNHNIYKLYGACLNIPV 100
 DB 1 VVGKAGFVVCWKAKRAKDVAIKQIESESRKAFIVELRQLSRVNHNIYKLYGACLNIPV 60
 QY 101 CLVMEYAEGLSYNVLHGAERPLPYRTAAHMSWCLQSGVAVLHMQPALLIHRDLKPP 160
 DB 61 CLVMEYAEGLSYNVLHGAERPLPYRTAAHMSWCLQSGVAVLHMQPALLIHRDLKPP 120
 QY 161 NLLVAGSTVLKICDFTGACDICTHMTNNKGSAAAMAPEVFEESNSSEKCDVFSWGIILW 220
 DB 121 NLLVAGSTVLKICDFTGACDICTHMTNNKGSAAAMAPEVFEESNSSEKCDVFSWGIILW 180
 QY 221 EVITRRKPFDEIGGPAFRIMAVHNGTRPPLIKNLPKPIESLMTRCMSKDPQSPSMEEI 280
 DB 181 EVITRRKPFDEIGGPAFRIMAVHNGTRPPLIKNLPKPIESLMTRCMSKDPQSPSMEEI 240
 QY 281 VKIMTHLMRYFPGADEPLQYPCQYSDGQNSATSTGSEFMDIASTNTSNKSDTNMEQVPA 340
 DB 241 VKIMTHLMRYFPGADEPLQYPCQYSDGQNSATSTGSEFMDIASTNTSNKSDTNMEQVPA 300
 QY 341 TNDTIKRLSKLLKNQAKQSEGRSLGASHGSSVESLPPTSEGRMSADMEIEARLA 400
 DB 301 TNDTIKRLSKLLKNQAKQSEGRSLGASHGSSVESLPPTSEGRMSADMEIEARLA 360
 QY 401 ATTNGQPRRRSIODLTVTGTEPGQVSSRSSSPSVRMITTSSEKPTRSHPTDDST 460
 DB 361 ATTNGQPRRRSIODLTVTGTEPGQVSSRSSSPSVRMITTSSEKPTRSHPTDDST 420
 QY 461 DTNGSDNSIPMAVYLLDHOLOPLAPCNSKESMAVFEQCKMAQSEYMKVQTEIALLQK 520
 DB 421 DTNGSDNSIPMAVYLLDHOLOPLAPCNSKESMAVFEQCKMAQSEYMKVQTEIALLQK 480
 QY 521 QELVAVELDDQEQQNTSRLVQEHKKLLDENKSLSTYYQCKKQLEVIRSQQKRGQTS 579
 DB 481 QELVAVELDDQEQQNTSRLVQEHKKLLDENKSLSTYYQCKKQLEVIRSQQKRGQTS 539

RESULT 3
 Q9NTR2 PRELIMINARY; PRT; 566 AA.
 ID Q9NTR2;
 AC Q9NTR2;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE D154G14.1.2 (Mitogen-activated protein kinase kinase 7
 (TGF-beta activated kinase 1b (TAK1))) (Fragment).
 GN MAP3K7.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tracey A.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 CC -i- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AL121964; CAB87606.1; -.
 DR HSSP; P12931; 1FMK.
 DR InterPro; IPR000719; Euk_pkinase.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR InterPro; IPR004040; STY_pkinase.
 DR InterPro; IPR01245; Tyr_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Euk_pkinase; 1.
 DR SMART; SM00221; STYKc_1
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

DR PROSITE; PS00101; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW Kinase; Serine/threonine-protein kinase.
FT NON TER 1
SQ SEQUENCE 566 AA; 62972 MW; 32CDAC1211B200CF CRC64;

Query Match 92.8%; Score 2795.5; DB 4; Length 566;
Best Local Similarity 95.1%; Pred. No. 5.8e-182;
Matches 536; Conservative 0; Mismatches 1; Indels 27; Gaps 1;

QY 41 VVGRGAFVWCKAKWRAKDVAKQIESERKAFIVELROLSRVNHPIVVKLYGACLNVP 100
Db 1 VVGRGAFVWCKAKWRAKDVAKQIESERKAFIVELROLSRVNHPIVVKLYGACLNVP 60

QY 101 CLVMEYAEGLSYLVNHLHGAELPYTTAAHMSWCLQSQGVAYLHSMQPKALIHRLDKPP 160
Db 61 CLVMEYAEGLSYLVNHLHGAELPYTTAAHMSWCLQSQGVAYLHSMQPKALIHRLDKPP 120

QY 161 NLLLVAGGTVLKICDFGTACDIQTHMTNNKGSAAWMAPEVFEKCDVFSWGIIW 220
Db 121 NLLLVAGGTVLKICDFGTACDIQTHMTNNKGSAAWMAPEVFEKCDVFSWGIIW 180

QY 221 EVITRRKPFDEIGGPAFRIMWAVHNGRPPLIKNLKPIESLMTRCWSKDPSPRPSMEEI 280
Db 181 EVITRRKPFDEIGGPAFRIMWAVHNGRPPLIKNLKPIESLMTRCWSKDPSPRPSMEEI 240

QY 281 VKIMTHLMRYPPGGADEPLQPCQYSDSQSNATSTGSMFMDIASTNTSNKSDTNMQVPA 340
Db 241 VKIMTHLMRYPPGGADEPLQPCQYSDSQSNATSTGSMFMDIASTNTSNKSDTNMQVPA 300

QY 341 TNDTIKLESKLLKNQAQKQSEGRSLGSHGSSVESLPPTSEKGRMSADMSEIEARIA 400
Db 301 TNDTIKLESKLLKNQAQKQSEGRSLGSHGSSVESLPPTSEKGRMSADMSEIEARIA 360

QY 401 ATT-----GNGQPRRSIQDLTVTGTETGQVSSRSSSP 433
Db 361 ATTAYSFKPKGHRKTASFGNILDVPEIVISGNGQPRRSIQDLTVTGTETGQVSSRSSSP 420

QY 434 SVRMITTSPTSEKPTRSHPTDDSDTNGSDNSIPMAYLTLDHQLQPLAPCPNSKESM 493
Db 421 SVRMITTSPTSEKPTRSHPTDDSDTNGSDNSIPMAYLTLDHQLQPLAPCPNSKESM 480

QY 494 AVPEHQCKMAQYKMKVQTEIALLQKQELVAELDQDEKQDQNTSLRVQEHKKLLDENKS 553
Db 481 AVPEHQCKMAQYKMKVQTEIALLQKQELVAELDQDEKQDQNTSLRVQEHKKLLDENKS 540

QY 554 LSTYVQCKKQLEVIRSQQKRGTS 579
Db 541 LSTYVQCKKQLEVIRSQQKRGTS 566

RESULT 4
O73613
ID O73613 PRELIMINARY; PRT; 616 AA.
AC O73613;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-WAR-2002 (TrEMBLrel. 20, Last annotation update)
DE TAK1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98130593; PubMed=9463380;
RA Shibuya H., Iwata H., Masuyama N., Gotoh Y., Yamaguchi K., Irie K.,
RA Matsumoto K., Nishida E., Ueno N.;
RT "Role of TAK1 and TAB1 in BMP signaling in early Xenopus
RT development.";
RL EMBO J. 17:1019-1028(1998).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

DR EMBL; U92030; AAC14008.1; -.
DR HSSP; P12931; 1FMK.
DR InterPro; IPR000719; Euk_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR004040; STY_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_kinase; 1.
DR SMART; SM00221; STYKc; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00108; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW ATP-binding; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 616 AA; 68464 MW; 493AD2A05ADC38B6 CRC64;

Query Match 87.5%; Score 2638.5; DB 13; Length 616;
Best Local Similarity 84.3%; Pred. No. 3.1e-171;
Matches 519; Conservative 20; Mismatches 28; Indels 49; Gaps 7;

QY 12 SSSAGEMIEAFSQVLNFEIDYKEIEVEEVVGRGAFVWCKAKWRAKDVAKQIESER 71
Db 2 SATSAEMIETP-PVLNFEIDYKEIEVEEVVGRGTFGVVCKAKWRAKDVAKQIESER 60

QY 72 KAFIVELQRLSRVNHPIVVKLYGACLNVPCLVMEYAEGLSYLVNHLHGAELPYTTAAHAM 131
Db 61 KAFIVELQRLSRVNHPIVVKLYGACLNVPCLVMEYAEGLSYLVNHLHGAELPYTTAAHAM 120

QY 132 SWCLQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKICDFGTACDIQTHMTNNKG 191
Db 121 SWCLQCAQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKICDFGTACDIQTHMTNNKG 180

QY 192 SAAMWAPVFEKGSYSEKCDVFSWGIILWEVITRRKPKPDEIGGPAFRIMWAVHNGTRPPL 251
Db 181 SAAMWAPVFEKGSYSEKCDVFSWGIILWEVITRRKPKPDEIGGPAFRIMWAVHNGTRPPL 240

QY 252 IKNLKPKIESLMTRCWSKDPSPRPSMEIIVKIMTHLMRYPPGGADEPLQPCQYSDSQSN 311
Db 241 IKNLKPKIESLMTRCWSKDPSPRPSMEIIVKIMTHLMRYPPGGADEPLQPCQYSDSQSN 300

QY 312 SATSTGSMFMDIASTNTSNKSDTNME---QVPATNDTIKLESK---LKNQAQKQSESG 364
Db 301 SATSTGSCIDITSTNTSNKSDIHIPEQDFQASANDTIKIESKLAQHLKNQAQKQSESG 360

QY 365 RLSLGASHGSSVESI-----PP-----TSEKGRMSADMSEIEARIAAT----- 403
Db 361 RLSLPPSRGSSVESISEIRGPPPTLTGTSEKGRMSADMSEIEARISASTAVVKPRGHRK 420

QY 404 -----GNGQPRRSIQDLTVTGTETGQV---SSRSSSPSVRMITTSGP 443
Db 421 TASFGNILDVPEIITAGNGQQRSSVQDLTVGTESQESRNSRSSSPSVRMITTSGP 480

QY 444 TSEKPTRSHPTDDSDTNGSDNSIPMAYLTLDHQLQPLAPCPNSKESMAVFEHQCKMA 503
Db 481 TDPKPRGLPWAPDESDTNGSDNSIPMAYLTLDHQLQPLAPCPNSKESMAVFEHQCKMA 540

QY 504 QEYMKVQTEIALLQKQELVAELDQDEKQDQNTSLRVQEHKKLLDENKSLSTYYQCKK 563
Db 541 QEYMKVQTEIALLQKQELVAELDQDEKQDQNTSLRVQEHKKLLDENKSLSTYYQCKK 600

QY 564 QLEVIRSQQKRGTS 579
Db 601 QLEVIRSQQKRGTS 616

RESULT 5
Q9NZ70
ID Q9NZ70 PRELIMINARY; PRT; 491 AA.
AC Q9NZ70; Q9NTR3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE TGF beta-activated kinase splice variant d (DJ154G1.1.4)
DE (mitogen-activated protein kinase kinase 7 (TGF-beta activated
DE kinase 1d (TAK1))).

CN TAK1 OR MAP3K7.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=2056765; PubMed=1118615;
 RA Demsey C.E., Sakurai H., Sugita T., Guesdon F.;
 RT "Alternative splicing and gene structure of the transforming growth
 RT factor beta-activated kinase 1.";
 RT Biochim. Biophys. Acta 1517:46-52(2000).
 RN [2]
 RP SEQUENCE OF 41-491 FROM N.A.
 RA Tracey A.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AF218074; AAF27652.1; -;
 DR EMBL; AL121964; CAB87605.1; -;
 DR HSSP; P08631; 1AD5.
 DR HSSP; P12931; 1FMK.
 DR InterPro; IPR000719; Euk_pkinase.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR InterPro; IPR004040; STY_pkinase.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Euk_pkinase; 1.
 DR SMART; SM00221; STYKC; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 KW SEQUENCE 491 AA; 53739 MW; B7D8832E286A59C5 CRC64;

Query Match 83.3%; Score 2512; DB 4; Length 491;
 Best Local Similarity 99.0%; Pred. No. 9e-163; Indels 0; Gaps 0;
 Matches 481; Conservative 0; Mismatches 5;

QY 1 MSTASASASSSSAGMIFAPSOVLNFEIDYKEIEVEEVVGRGAFVCAKAKMRKADV 60
 DB 1 MSTASASASSSSAGMIFAPSOVLNFEIDYKEIEVEEVVGRGAFVCAKAKMRKADV 60
 QY 61 AIKQIESESEKKA FIVELRQLSRVNHPIVLYGACINPVCLMEYAEAGSLYNVLHGAE 120
 DB 61 AIKQIESESEKKA FIVELRQLSRVNHPIVLYGACINPVCLMEYAEAGSLYNVLHGAE 120
 QY 121 PLPYTTAAHMSWCLQCSQGVAYLHSMQKALIHRLDKPNNLLVAGGTVLKICDEGTAC 180
 DB 121 PLPYTTAAHMSWCLQCSQGVAYLHSMQKALIHRLDKPNNLLVAGGTVLKICDEGTAC 180
 QY 121 PLPYTTAAHMSWCLQCSQGVAYLHSMQKALIHRLDKPNNLLVAGGTVLKICDEGTAC 180
 DB 121 PLPYTTAAHMSWCLQCSQGVAYLHSMQKALIHRLDKPNNLLVAGGTVLKICDEGTAC 180
 QY 181 DIQTHMTNNKGSAAAMAPAEVFECSNYSKCDVFSWGIILMEVITRRKKPDEIGAPFRIM 240
 DB 181 DIQTHMTNNKGSAAAMAPAEVFECSNYSKCDVFSWGIILMEVITRRKKPDEIGAPFRIM 240
 QY 241 WAVHNGTRPPLIKNLPKPIESLMTKRCWKDPSQPSMEEIVKIMTHLMRYFFGADDEPQY 300
 DB 241 WAVHNGTRPPLIKNLPKPIESLMTKRCWKDPSQPSMEEIVKIMTHLMRYFFGADDEPQY 300
 QY 301 PCQYSDGQSNATSTGSPMDIASTNTSNKSDTNMEQVPAINTDTIKRLSKLKNQAKQO 360
 DB 301 PCQYSDGQSNATSTGSPMDIASTNTSNKSDTNMEQVPAINTDTIKRLSKLKNQAKQO 360
 QY 301 PCQYSDGQSNATSTGSPMDIASTNTSNKSDTNMEQVPAINTDTIKRLSKLKNQAKQO 360
 DB 301 PCQYSDGQSNATSTGSPMDIASTNTSNKSDTNMEQVPAINTDTIKRLSKLKNQAKQO 360
 QY 361 SESGRLLGASHGSSVSLPPTSEKGMASDMSEIEARIAATTGNGQPRRRSIQDLTVTG 420
 DB 361 SESGRLLGASHGSSVSLPPTSEKGMASDMSEIEARIAATTGNGQPRRRSIQDLTVTG 420
 QY 421 TEPQVSSRSSSPSVRMITTSSEKPTRSHPTPDSTDTNGSDNSIPMAVLTLDHQL 480
 DB 421 TEPQVSSRSSSPSVRMITTSSEKPTRSHPTPDSTDTNGSDNSIPMAVLTLDHQL 480
 QY 481 QPLAPC 486

DB 481 QARTSC 486

RESULT 6
 Q9NTR4 PRELIMINARY; PRT; 478 AA.
 AC Q9NTR4;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE D154G14.1.3 (Mitogen-activated protein kinase kinase 7
 DE (TGF-beta activated kinase 1c (TAK1))) (Fragment).
 CN MAP3K7.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tracey A.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AL121964; CAB87604.1; -;
 DR HSSP; P12931; 1FMK.
 DR InterPro; IPR000719; Euk_pkinase.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR InterPro; IPR004040; STY_pkinase.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Euk_pkinase; 1.
 DR SMART; SM00221; STYKC; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR Kinase; Serine/threonine-protein kinase.
 FT NOW TRER 1
 SO SEQUENCE 478 AA; 52482 MW; 177CC8CBA8D8DF8 CRC64;

Query Match 76.4%; Score 2301.5; DB 4; Length 478;
 Best Local Similarity 93.2%; Pred. No. 1.8e-148;
 Matches 441; Conservative 0; Mismatches 5; Indels 27; Gaps 1;

QY 41 VVGGAEGVVCCKAKRKAQDAIKQIESESEKKA FIVELRQLSRVNHPIVLYGACINPV 100
 DB 1 VVGGAEGVVCCKAKRKAQDAIKQIESESEKKA FIVELRQLSRVNHPIVLYGACINPV 100
 QY 101 CLVMEYAEAGSLYNVLHGAEPLPYTTAAHMSWCLQCSQGVAYLHSMQKALIHRLDKP 160
 DB 101 CLVMEYAEAGSLYNVLHGAEPLPYTTAAHMSWCLQCSQGVAYLHSMQKALIHRLDKP 160
 QY 61 CLVMEYAEAGSLYNVLHGAEPLPYTTAAHMSWCLQCSQGVAYLHSMQKALIHRLDKP 120
 DB 61 CLVMEYAEAGSLYNVLHGAEPLPYTTAAHMSWCLQCSQGVAYLHSMQKALIHRLDKP 120
 QY 161 NLLVAGGTVLKICDEGTACDIQTHMTNNKGSAAAMAPAEVFECSNYSKCDVFSWGIILW 220
 DB 161 NLLVAGGTVLKICDEGTACDIQTHMTNNKGSAAAMAPAEVFECSNYSKCDVFSWGIILW 220
 QY 221 EVITRRKPFDEIGAPFRIMAVHNGTRPPLIKNLPKPIESLMTKRCWKDPSQPSMEEI 280
 DB 221 EVITRRKPFDEIGAPFRIMAVHNGTRPPLIKNLPKPIESLMTKRCWKDPSQPSMEEI 280
 QY 281 VKIMTHLMRYFFGADDEPQYPCQYSDGQSNATSTGSPMDIASTNTSNKSDTNMEQVPA 340
 DB 281 VKIMTHLMRYFFGADDEPQYPCQYSDGQSNATSTGSPMDIASTNTSNKSDTNMEQVPA 340
 QY 341 TNDTIRKLESLKLNQAKQOESGRSLGASHGSSVSLPPTSEKGMASDMSEIEARIA 400
 DB 341 TNDTIRKLESLKLNQAKQOESGRSLGASHGSSVSLPPTSEKGMASDMSEIEARIA 400
 QY 401 ATT-----GNGQPRRRSIQDLTVTGTEPQVSSRSSSP 433
 DB 361 ATTAVSKRGRKRTASGNILVPEIYISGNGQPRRRSIQDLTVTGTEPQVSSRSSSP 420
 QY 434 SVRMITTSSEKPTRSHPTPDSTDTNGSDNSIPMAVLTLDHQLQPLAPC 486

Db 421 SVRMITTSKPTRSRHPWTPDDSTDTNGSDNSIPMAYLTLDHQLOARTSC 473

RESULT 7

Q9UG54 PRELIMINARY; PRT; 260 AA.
AC Q9UG54;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE Hypothetical 28.9 kDa protein.
GN DKF2P586F0420.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=UTERUS;
RC Wambutt R., Heubner D., Mewes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL050393; CAB43687.2; -;
KW Hypothetical protein.
SQ SEQUENCE 260 AA; 28870 MW; D24F292932FB7F72 CRC64;

Query Match 43.5%; Score 1310; DB 4; Length 260;
Best Local Similarity 99.6%; Pred. No. 2e-81;
Matches 259; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 320 MDIASTNTSNKSDTNNEQVPATNDTIKRLSKLLKNQAKQSGESRLSLGASHGSSVESL 379
|||||

Db 1 MDIASTNTSNKSDTNNEQVPATNDTIKRLSKLLKNQAKQSGESRLSLGASHGSSVESL 60
|||||

QY 380 PPTSEKGRMSADNSETIARTAATNGCQPRRSIQDLTVTGTGPGVSSRSSPSVRMIT 439
|||||

Db 61 PPTSEKGRMSADNSETIARTAATNGCQPRRSIQDLTVTGTGPGVSSRSSPSVRMIT 120
|||||

QY 440 TSQPTSEKPTRSRHPWTPDDSTDTNGSDNSIPMAYLTLDHQLOPLAFCPNKSKESMAVFEGH 499
|||||

Db 121 TSQPTSEKPTRSRHPWTPDDSTDTNGSDNSIPMAYLTLDHQLOPLAFCPNKSKESMAVFEGH 180
|||||

QY 500 CKMAQYMKVQTEIALLQKQBELVAELDQDEKQNTSLRQVHKLLDENKSLSTYYQ 559
|||||

Db 181 CKMAQYMKVQTEIALLQKQBELVAELDQDEKQNTSLRQVHKLLDENKSLSTYYQ 240
|||||

QY 560 QCKKQLEVIRSQQKRGQTS 579
|||||

Db 241 QCKKQLEVIRSQQKRGQTS 260
|||||

RESULT 8

Q9V3Q6 PRELIMINARY; PRT; 678 AA.
AC Q9V3Q6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Putative TAK1 protein (CG1388 protein) (LD42274P).
GN TAK1 OR CG1388 OR CG18492.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Cocayne J.D.,
RA Ananides F.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.P., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fesler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush P., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Morkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Takatsu Y., Nakamura M., Stapleton M., Danos M., Matsumoto M.,
RA O'Connor M.B., Shibuya H., Ueno N.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AE003571; AAF50895.1; -;
DR EMBL; AF199466; AAF06815.1; -;
DR EMBL; AY051953; AAK93377.1; -;
DR HSSP; P08631; IAD5
DR Flybase; FBgn026323; Tak1.
DR InterPro; IPR00719; Euk.pkinase.
DR InterPro; IPR001990; Granin.
DR InterPro; IPR002290; Ser thr.pkinase.
DR InterPro; IPR004040; STY.pkinase.
DR InterPro; IPR001245; Tyr.pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Euk.pkinase; 1.
DR SMART; SM00221; STYKc; 1.
DR PROSITE; PS00422; GRANINS 1; UNKNOWN 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Hypothetical protein; Kinase;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 678 AA; 75674 MW; 87EBA80CDB8CDE45 CRC64;

Query Match 32.1%; Score 966; DB 5; Length 678;

Best Local Similarity 36.0%; Pred. No. 1.8e-57;

Matches 226; Conservative 97; Mismatches 207; Indels 98; Gaps 14;

```
QY 18 MIEAPSVLNFEEIDYKEIEVEEVNGARGVYCKAKWRADVAIKQISESEKRAPIVE 77
DB 1 MATASLDLAAAYVDFEILREKVGSGYVCKAWRKLAIVEKFAASAKOIKE 60
QY 78 LROLSRVNHNPIVLYG--ACLNPVCLMVEAEGSLVYVHGAEPYPYTAHANSWCL 135
DB 61 VKOLSRVKNHNIILAHIOISSYQATYILMFAGGSLAHNLHG-KYKPAVSLAHANSAR 119
QY 136 QCSQGVAYLHSPKALIHBDLKPILLVAGTVLKICDFGACDICTHMTNNKSAAM 195
DB 120 QCAGEGLAYLHAMPKPIHRDVKPMLLLTNKRNLIKICFGTVAADKSTMTNNKSAAM 179
QY 196 MAEVEPVGSGYSEKCDVFSWGIILMEVITRRKPFDELGGAFRIMAVNGRRPILIKL 255
DB 180 MAEVEPVGSGYSEKCDVFSWGIILMEVITRRKPFDELGGAFRIMAVNGRRPILIKL 238
QY 256 PKDIESIMTRCWSKDSQSPSMEEIVKIMTHLRYPFGADEPQY----- 300
DB 239 PKRIEDLMTCKMTVPEDRPSMQYIVGVHMEIVKDYTGADKALETFVNOQIVTKESDGT 298
QY 301 -----PCQYSEGOQN-----SATSTGSEMDIASNTSNKSDTN 334
DB 299 VAAQPSLSSQEGELSPSSSTQLTPTTAANVANVAIAISKTTTSMTEINTSSSDITPTN 358
QY 335 MEQVP-----ATN--DTIKLESKLKN-----QAKQSSGRLSGASHGSSV 376
DB 359 SGOLDNPLFYVYVYVNRMDALPEESNESRSDSNLTSSAEATQRLTIRNGMILWACKPM 418
QY 377 ESLEPTSEGRMSADMSELARIATATGNQPRRSIODLTVTGTEP-----CQ 425
DB 419 EQL--TLVDANGPDLSPSSSSSSSTNAKSDGER---LTVTDTPFVMTTDLSDNNNG 472
QY 426 VSSRSSPSVRMTTGTGPTSEKTRSHPW-----TPDSDTNGSDNSIPMAVTL 476
DB 473 IHAHNSGLSHANGQARDELOEHEDEIYNSLDVDVDPEDENDEGTEGSIAE---IL 529
QY 477 DHLOLAPCPNKSMAVPECHCKMAOEYMKVOTEIALLQKQELVAVELDODEKQDN 536
DB 530 DEPLQEPPTPNDASQOLYIRDRHMAKEVLSYDTLVYAQDFKDLIVQMRTKEQK- 588
QY 537 TSLVQEHKKLDEKNSLSTYYQCKKQ 564
DB 589 -----QELLRKMKDKGLOSLVNNLQQ 611

RESULT 9
Q9VCVO PRELIMINARY; PRT; 252 AA.
AC Q9VCVO;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE CG4803 protein.
GN CG4803.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Ephydriota; Neoptera; Enopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydriidae; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champ M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abriti J.F., Abmayant A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
```

```
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotier P.,
RA Butris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Dudson K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischman W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Houtlin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jaisl M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merklov G., Mishina N.V., Mobary B., Murphy L., Murty D.M., Nelson D.L.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Murty D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson W., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner C., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
RA Glabs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: AEO03740; AAF56055.1;
DR FLYBase; FBgn0039015; CG4803.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase.1.
DR ProDom; PD000001; Euk_pkinase.1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP.1.
DR PROSITE; PSS0011; PROTEIN_KINASE_DOM.1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST.1.
DR ATP-binding; Serine/threonine-protein kinase; Transferase.
KW SEQUENCE 252 AA; 28853 MW; 8E006F2EAB68D17 CRC64;

Query Match 16%; Score 505; DB 5; Length 252;
Best Local Similarity 40.6%; Pred. No.1e-26;
Matches 99; Conservative 48; Mismatches 71; Indels 26; Gaps 4;

QY 29 EEIDYKEIEVEEVNGARGVYCKAKWRADVAIKQISESEKRAPIVELROLSRVNHNPN 88
DB 6 EGVEYEEIQTKELGTGFGYSYVRAVWRNNEILAKRIRECECEKXIRELYOLTASHVN 65
QY 89 IVKLYGACLPVC--LMEYAEGLSLVYVHGAEPYPYTAHANSWCLQCSQGVAYLHS 146
DB 66 IVELYGTSRHEGCLALIMEFVDGSLSSFLH-AKSPSYSHAHAFWMAHQIAGIAYLHG 124
QY 147 MQPKALIHBDLKPENLLVAGTVLKICDFGACDICTHMTNNKSGAAMAP----- 199
DB 125 MQPKAVIHRDILPNTLLCEKGLKLIKICFGTVAADKSTMTNNKSGAAMAPVRELDFD 184
QY 200 -----VEGSGYSEKCDVFSWGIILMEVITRRKPFDELGGAFRIMAVNGRRPILIKL 244
DB 185 KSNRIINOPTGFKVLQGNKPKPEKCDVYSMAITPFWELISRKKEPFQY-NTLPELMAIN 243
QY 245 NGTR 248
DB 244 EGKR 247

RESULT 10
Q9HCC4 PRELIMINARY; PRT; 455 AA.
AC Q9HCC4;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
```



```

Db 173 MADEVLIQSLPSECTDYSGVILWEMLTREVPFKGLEG--LQVAMLVYEKMERLTI PSS 230
QY 255 LKPIESLMTKRCWSKDSOPSPSMEBEIYKINTHLMRYPFGADEPLQYPCQYSDGOSNSKT 314
Db 231 CPSPFELLHQCEWAPAKKPSFKQIISLI-----ESMSNDT- 267
QY 315 STSPFMDIASTNTSNKSDTMEQVPATNDTIKRLSKLTKNQAKQOSESGRSL---GAS 371
Db 268 ---NLDPQCSPLHNAKEMNCE-IEATLERLKLERDLSEKQELKERERRLKMEQKLT 323
QY 372 HGSSVESLPPTS-----EGKMSADNSEIEARI-AAITGNGQPRRRSIQDITVTGT 421
Db 324 EQSNTPLLLPLLSARMSSESYFESKTEESNAGEMSQITAAISNGHGGMNPLQAMMLMGF 383
QY 422 EPGQVSSRSSSPSV 435
Db 384 --GDVFSMNKAGAV 395

```

RESULT 12

```

O9H2N5 PRELIMINARY; PRT; 1066 AA.

```

```

AC O9H2N5;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Mixed lineage kinase MLK1 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA McNea J.J., Dower S.K., Guesdon F.;
RT "cDNA sequence and gene organisation of mixed lineage kinase 1.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
DR EMBL; AF251442; AAG44591.1; -.
DR HSSP; P29355; ISEM.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001452; SH3.
DR pfam; PF00069; pkinase; 1.
DR PRINTS; PR00018; SH3; 1.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00326; SH3; 1.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50002; SH3; 1.
DR KINASE; SH3 domain.
FT NON_TER 1
SQ SEQUENCE 1066 AA; 118463 MW; EDD08EBE7482723 CRC64;

```

```

Query Match 15.3%; Score 462.5; DB 4; Length 1066;
Best Local Similarity 27.0%; Pred No. 5.4e-23;
Matches 141; Conservative 81; Mismatches 192; Indels 109; Gaps 15;

```

```

QY 3 TASAASSSSSSSAGE--MIEAPSOVLNFEIDYKEIEVEEVVGRGAFGVCAKAKRAKV 60
Db 60 TRSAFSSSCQGGEDPCYPIQL---EIDFAELTLEIIGIGFGKVVRAFIIGSV 116
QY 61 AIK--QIESESRKAFIVLRQ---LSRVNHPNIVKLYGACLANP--VCLVMEVAGGSL 112
Db 117 AVKARHDPDEDIDISQTIENVROEAKFLKAPNIIALRGVCLKEPNCVLVMEFARGGSL 176

```

```

QY 113 YNVILGAEPLPYTAAAHAMSWCLQCSQGVAVLHSMQPKLIHRDKLPNNLLV----- 165
Db 177 NRVLGKRIIPPD-----LVNVAVOIAGMNYIHDEALVPIIHRDLKSNLILIQKVENGD 232
QY 166 AGTVLKIICDFGTACD-IQTHMTNNKSAAMAWEVEGGSNSYSEKCDVFSWGIILEVIT 224
Db 233 LSNKILKITDGLAREWHRRITKMSAAGTYAMMADEVIRASMFSGSDVMSYGVILMELLT 292
QY 225 RRPKDEIGGAFIIMAVHNGTRPPLIKLKPRIEISLMTKRCWSKDSOPSPSMEBEIYKIM 284
Db 293 GEVPRGIDGLAVAVGVAM-NKALPIPTCPEPFAXIMEDCNPPHSPSPFTNIIIDQL 351
QY 285 THLNRYPGADEPLQYPCQYSDGOSNSATSTGSF-----MDIASNTSNKSDTMEQV 338
Db 352 T-----TIESGFEPKPSFHCLODNMKHEIQEMFDQL 385
QY 339 PATNDTIKRLSKLTKNQAKQOSESGRL-----S 367
Db 386 RAKEKEIETWEELTRAALQCKQOEELRRREQELAREIDILERELNIITIHQLCOEKPR 445
QY 368 LGASHGSSVESLPPTSFGKMSADMSIEARIATGNGQPRRRSIQDITVTGTGQVS 427
Db 446 VKRRKGKFRKSRRLKDKGNRISLP-SDQHKFTVOASPTWDKRKSLL-----IN 492
QY 428 SRSSSPSVRMITTSQPTSEKPTRSHPTPDSDTDTGSDNSIP 470
Db 493 SRSSSP-----ASPTIIRLRAIQLTPGESSKTKWGRSSVVP 528

```

RESULT 13

```

O9JUI5 PRELIMINARY; PRT; 850 AA.

```

```

AC O9JUI5;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Mixed lineage kinase 3.
GN MLK3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=129/Ola;
RX MEDLINE=20354997; PubMed=10894943;
RA Saridaki A., Ferraz C., Demaille J., Scherrer G., Roux A.-F.;
RT "Genomic sequencing reveals the structure of the Kmk6 and Map3k1
RT genes and their close vicinity to the Sipa1 gene on mouse chromosome
RT 19.";
RL Cytogenet. Cell Genet. 89:85-86(2000).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC EMBL; AF155142; AAF73281.1; -.
DR HSSP; P29355; ISEM.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001452; SH3.
DR pfam; PF00069; pkinase; 1.
DR pfam; PF00018; SH3; 1.
DR PRINTS; PR00452; SH3DOMAIN.
DR ProDom; PD000001; Euk_pkinase; 1.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00326; SH3; 1.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50002; SH3; 1.
DR ATP-binding; Kinase; SH3 domain; Serine/threonine-protein kinase;
KW Transferase.

```


RESULT 15
Q8VDG6 PRELIMINARY; PRT; 1001 AA.
ID Q8VDG6
AC Q8VDG6
DT 01-MAR-2002 (T-EMBLrel. 20, Created)
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE Similar to mitogen-activated protein kinase kinase 9.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC021891; AAH21891.1; -
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR001452; SH3.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR Pfam: PF00018; SH3; 1.
DR PRINTS: PRO0452; SH3DOMAIN.
DR ProDom: PD000001; Euk_pkinase; 1.
DR ProDom: PD000066; SH3; 1.
DR SMART: SM00326; SH3; 1.
DR SMART: SM00220; S_TKc; 1.
DR SMART: SM00219; TyKc; 1.
DR PROSITE: PS00107; PROTEIN KINASE ATP; UNKNOWN_1.
DR PROSITE: PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.
DR PROSITE: PS00002; SH3; 1.
KW kinase.
SQ SEQUENCE 1001 AA; 109983 MW; E10042C68B9953C CRC64;
Query Match 15.2%; Score 457; DB 11; Length 1001;
Best Local Similarity 26.7%; Pred. No. 1.2e-22;
Matches 159; Conservative 89; Mismatches 219; Indels 128; Gaps 23;
QY 31 IDYKEIEVEEVGGAFGVVCYAKWRAKDVAIKQIESSEKAFIV-----ELRQLSRV 84
DB 105 VDEPERLEKELIGAGGFGQVYRAITWQGEVAVAAARRDEQDAAAABSVREARLFAWL 164
QY 85 NHPNIVKLYGACLNP--VCLVMEYAEAGSLYNVLHGAEPPLPYTTAAH-----AMSW 133
DB 165 RHENITQLRGVCLRQPHLCVLEFARGALNRALAAASDPRAAPRRARIPQVLYVMW 224
QY 134 CLQCSQGVAYLHSMOKALIHRLKPPNLLVAG-----GTVLKICDFTACD-IQTH 185
DB 225 AVQIARGMLYLHEEAVVPIHRLDKSSNILLLEKIEHDDICNKTLEITDFTGLAREWHRTT 284
QY 186 MTNNKSSAAMMAPEVEFGSNVSEKCVFSGWIIIMEVITRRKPFDEIGGPAFRIMAVAHN 245
DB 285 RMSAAGTYAMMAPEVIRSSIFSGSDIMSYGVLLMELTGEVPRGIDGLAAYGVAVVK 344
QY 246 GTRPPLIKULPKPIESLMTRCWSKDSQPSMEIVKIMTHL-----MRYFPGADE 296
DB 345 LTL-PIPTCTPEPFATLMEKCEWQDPIHPSFALLIQQLTAIEAVLTNMPQESFHSQE 403
QY 297 PLOYPQYSDSEGQSNATSTGSGFMDIASTNTSKSD---TNMEQVPATN--DTIKRLS 350
DB 404 DWLLEIQLMSELTKEKELRSREELSRALQOKSQELLRRREQOLAREIDVLER-EL 462
QY 351 KLLKNQAKOOS-----ESGRSLGSHSHSVES-----LPPTSEGRMSAD 391
DB 463 NVLIFQLSQEAPVKKRKGFRGRRLRLKDGHRISLPSPQHKITYQASPTLDKRSS-- 520
QY 392 MSEIEARIAATGNG--QPRRSIQDLVTGTPEPGQVSSRSSSPSVRM----- 438
DB 521 ---DSGLCSPPGSPFLMLPLRLAIQ---LTPSENNKTRGRN-----WVFRQEDFEDVXR 567

QY 439 -----TTSGP-----TSEKP-----TRSHPWTPDDSTDTNGSDNSIPMAVLTLDH 478
DB 568 SFKKKGCTWGPSSVQTKERPEGRERVRPLSDGNSPW-----SLLIKSQKTTPLASLFVDQ 623
QY 479 QLOPLAPCPNKEKMAVE---QCKMAQEMKVKQTEIALL-----QRQBELVAE 526
DB 624 -----PGSCBQKLVPEGLEHRRKPKQTKFPQAHVGLPLCKDSQREDSSEAE 670

Search completed: December 9, 2002, 22:58:22
Job time : 67.5512 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 9, 2002, 22:48:29 ; Search time 24.5928 Seconds
(without alignments)
692.718 Million cell updates/sec

Title: US-09-830-144-2
Perfect score: 3014
Sequence: 1 MSTASAASSSSSSAGEMIE.....QCKKOLEVIRSQQKROGTS 579

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5A COMB.pep:*
2: /cgn2_6/prodata/1/iaa/5B COMB.pep:*
3: /cgn2_6/prodata/1/iaa/6A COMB.pep:*
4: /cgn2_6/prodata/1/iaa/6B COMB.pep:*
5: /cgn2_6/prodata/1/iaa/PCTUS COMB.pep:*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3014	100.0	579	4	US-09-529-279-4
2	3014	100.0	590	4	US-09-529-279-15
3	481	16.0	455	3	US-09-221-235-5
4	481	16.0	455	3	US-09-221-928-5
5	481	16.0	455	4	US-09-221-527-5
6	481	16.0	455	4	US-09-221-236-5
7	481	16.0	455	4	US-09-221-416-5
8	481	16.0	455	4	US-09-221-245-5
9	481	16.0	455	4	US-09-163-115-5
10	481	16.0	455	4	US-09-221-528-5
11	481	16.0	455	4	US-09-593-553-5
12	481	16.0	455	4	US-09-221-237-5
13	418	13.9	835	4	US-09-291-839-2
14	412	13.7	668	1	US-08-205-018-2
15	412	13.7	859	1	US-08-395-580-2
16	412	13.7	859	5	PCT-US95-02792-2
17	398	13.2	821	1	US-07-928-464-2
18	398	13.2	821	1	US-08-003-311B-2
19	398	13.2	821	1	US-08-261-432-2
20	398	13.2	821	5	PCT-US93-07347-2
21	393	13.0	263	3	US-09-035-706-5
22	393	13.0	263	3	US-08-955-841-5
23	393	13.0	263	4	US-09-390-425-5
24	393	13.0	263	4	US-09-566-906-5
25	390.5	13.0	787	4	US-09-188-930-334
26	380.5	12.6	536	4	US-09-188-930-185
27	379.5	12.6	1584	4	US-09-457-040B-27

28	363.5	12.1	527	4	US-08-426-509A-10	Sequence 10, Appl
29	363.5	12.1	527	5	PCT-US95-05008-10	Sequence 10, Appl
30	358.5	11.9	620	4	US-08-426-509A-9	Sequence 9, Appl
31	358.5	11.9	620	5	PCT-US95-05008-9	Sequence 9, Appl
32	356.5	11.8	1052	3	US-08-863-118-2	Sequence 2, Appl
33	356	11.8	625	1	US-08-391-615-3	Sequence 3, Appl
34	355.5	11.8	983	1	US-08-167-919A-10	Sequence 10, Appl
35	355.5	11.8	983	3	US-08-715-106-10	Sequence 10, Appl
36	354.5	11.8	983	2	US-08-449-645A-21	Sequence 21, Appl
37	354.5	11.8	983	2	US-08-702-367A-21	Sequence 21, Appl
38	354.5	11.8	983	5	PCT-US95-04681-21	Sequence 21, Appl
39	351.5	11.7	269	2	US-07-857-224B-79	Sequence 79, Appl
40	349.5	11.6	304	2	US-08-701-191A-27	Sequence 27, Appl
41	349.5	11.6	1052	3	US-08-863-118-1	Sequence 1, Appl
42	349.5	11.6	1052	4	US-09-377-310-2	Sequence 2, Appl
43	346.5	11.5	982	2	US-08-673-789-4	Sequence 4, Appl
44	345.5	11.5	680	1	US-08-542-363-4	Sequence 4, Appl
45	345.5	11.5	680	4	US-09-100-089-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-09-529-279-4
; Sequence 4, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529, 279
; CURRENT FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-279-4

Query Match 100.0%; Score 3014; DB 4; Length 579;
Best Local Similarity 100.0%; Pred. No. 3.8e-222;
Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MSTASAASSSSSSAGEMIEAPSOVLNFEEDIDYKEIEVEEVVGRGAFGVCKAKRAKV	60
Db	1	MSTASAASSSSSSAGEMIEAPSOVLNFEEDIDYKEIEVEEVVGRGAFGVCKAKRAKV	60
Qy	61	AIKQIESERKAFIVELRQLSRVNHNIKLYGACLNPCVCLMVEYAEGLVNLHGAE	120
Db	61	AIKQIESERKAFIVELRQLSRVNHNIKLYGACLNPCVCLMVEYAEGLVNLHGAE	120
Qy	121	PLPYTTAAHMSWCLQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGTGVLCIDFGTAC	180
Db	121	PLPYTTAAHMSWCLQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGTGVLCIDFGTAC	180
Qy	181	DIQTHMTNNKSAAMWAPPEFEGSNYSKCDVFSWGIIILWEVITRRKPPFDEIGGPAFRIM	240
Db	181	DIQTHMTNNKSAAMWAPPEFEGSNYSKCDVFSWGIIILWEVITRRKPPFDEIGGPAFRIM	240
Qy	241	WAVHNGTRPPLIKNLPKPIESLMTRCWSKDPSPORMEEIVKIMTHLMRYFPGADEPQY	300
Db	241	WAVHNGTRPPLIKNLPKPIESLMTRCWSKDPSPORMEEIVKIMTHLMRYFPGADEPQY	300
Qy	301	PCQYSDGQSNATSFTGSMFDIASTNTSNKSDTNMEQVATNDTIKLESKLLKNQAKQ	360

```
Db 301 PCQYSDGSGNSATSTGSPMDIASINTSNKSDTNMEQVPAINDTIKRLSKLKNQAKQ 360
Qy 361 SESGRSLGASHGSSVSESLPPTSEGRKMSADMEIEARIAATTGNGQPRRSIODLTVTG 420
Db 361 SESGRSLGASHGSSVSESLPPTSEGRKMSADMEIEARIAATTGNGQPRRSIODLTVTG 420
Qy 421 TEPGVSSRSSSPSVRMITTSGETSEKPTRSHPWTPDDSTDTNGSDNSIPMAVLTLDHQL 480
Db 421 TEPGVSSRSSSPSVRMITTSGETSEKPTRSHPWTPDDSTDTNGSDNSIPMAVLTLDHQL 480
Qy 481 QPLAPCPNKSESMAVFEQHCXMAOEYMKVOTELALLQKQELVAELDDKQONTSL 540
Db 481 QPLAPCPNKSESMAVFEQHCXMAOEYMKVOTELALLQKQELVAELDDKQONTSL 540
Qy 541 VOEHKKLIDENKSLSTYYQCKKQLEVIRSQQKRGOTS 579
Db 541 VOEHKKLIDENKSLSTYYQCKKQLEVIRSQQKRGOTS 579

RESULT 2
US-09-529-279-15
; Sequence 15, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529,279
; CURRENT FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 590
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-279-15

Query Match 100.0%; Score 3014; DB 4; Length 590;
Best Local Similarity 100.0%; Pred. No. 3.9e-222;
Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Db 361 SESGRSLGASHGSSVSESLPPTSEGRKMSADMEIEARIAATTGNGQPRRSIODLTVTG 420
Qy 421 TEPGVSSRSSSPSVRMITTSGETSEKPTRSHPWTPDDSTDTNGSDNSIPMAVLTLDHQL 480
Db 421 TEPGVSSRSSSPSVRMITTSGETSEKPTRSHPWTPDDSTDTNGSDNSIPMAVLTLDHQL 480
Qy 481 QPLAPCPNKSESMAVFEQHCXMAOEYMKVOTELALLQKQELVAELDDKQONTSL 540
Db 481 QPLAPCPNKSESMAVFEQHCXMAOEYMKVOTELALLQKQELVAELDDKQONTSL 540
Qy 541 VOEHKKLIDENKSLSTYYQCKKQLEVIRSQQKRGOTS 579
Db 541 VOEHKKLIDENKSLSTYYQCKKQLEVIRSQQKRGOTS 579

RESULT 3
US-09-221-235-5
; Sequence 5, Application US/09221235
; Patent No. 6043040
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/221,235
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-235-5

Query Match 16.0%; Score 481; DB 3; Length 455;
Best Local Similarity 30.4%; Pred. No. 6.2e-29;
Matches 132; Conservative 81; Mismatches 151; Indels 70; Gaps 18;

Qy 27 NFEIDYKEIEVEBVGRGAGVVCRAK--RAKDAIK--QIESSEKKAFLVELROL 81
Db 7 SFVQKFDLDLFFENCGGSGFSGYRAKMIQDKVAVKLLKIEKEAF-----IL 57

Qy 82 SRVNHPIVVKLYGACLP--VCLVMEYABEGSLYNVLHG--APLPIYTTAAHMSWCLQC 137
Db 58 SVLSHRNIIQFYGVILPEPNYGIYTEYASLSGLYDINSNRSEB--DMDHIMTATDV 114

Qy 138 SOGVAYLHSMQKALJHRDLKPNLLVAGTVLKICDFTACDIQTHMTNNK--GSAAM 195
Db 115 AKGMHYLLMEAPVAVYIHRDLKSRNVVIAADG-VLKICDFG-ASRFNHHTHMSLVGTFFW 172

Qy 196 MAPEVFECSNYSEKCDVFSWGIILMEVITRRKPEDEIGGPAFRIM-AYANGTRPPLIKN 254
Db 173 MAPEVIGSLPVSERCDTVSYGVVLMEMLTREVPFKGEG--LQVAVLVEKNERLTPSS 230

Qy 255 LKPIESIMTSCMSKDSQPSMEIYKIMTHLMRYFGADEPLQYCYQSDGQNSAT 314
Db 231 CPRSFABELHOCWEADAKKRPSPFOIISIL-----ESMSNDT- 267

Qy 315 STGSPMDIASINTSNKSDTNMEQVPAINDTIKRLSKLKNQAKQSESGRLSL--GAS 371
Db 268 ---SLPDKCNSFLHNKAKWRCF-IEATLERUKLEBRLSFEQGLKERERLKNMEQKLT 323

Qy 372 HGSVSESLPPTS-----EGKMSADMEIEARIAATT-GNGQPRRSIODLTVTG 421
Db 324 EGSNTPULLPLAAMSESYFESTBESNSAEMSCQITATSNGBGHWNSLOAMLMGF 383

Qy 422 EPGQVSSRSSSPSV 435
Db 384 --GDIFSMKAGAV 395

RESULT 4
```



```
US-09-221-928-5
; Sequence 5, Application US/09221928
; Patent No. 6121030
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,928
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-928-5

Query Match      16.0%; Score 481; DB 3; Length 455;
Best Local Similarity 30.4%; Pred. No. 6.2e-29;
Matches 132; Conservative 81; Mismatches 151; Indels 70; Gaps 18;

QY 27 NFEEDYKEIEVEEVGAGFVVCCKAW--RAKDVAIK--QIESERKAFIVELROL 81
Db 7 SFVQIKFDDLQFFENC GGSGFSGVYRAKWSQDKEVAVKLLKIEKEAE-----IL 57

QY 82 SRVNHPIVKLYGACLNLP--VCLVMEYAEAGSLYNVLHG--AEPLPYVYTAHAHMSWCLOC 137
Db 58 SVLSHRNIIQFYGVILEPPNYGIVTEYASLGSLDYINSRSEEM---DMDHIMTWATDV 114

QY 138 SQGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFGTACDIQTHMTNNK--GSAAW 195
Db 115 AKGMHYLHMEAPVKVIHRDLKSRNVVIAADG-VLKICDFG-ASRPHNHTTHMSLVGTFFW 172

QY 196 MAPEVFEGSNYSEKCDVFSWGIIILWEVITRRKPFDEIGGPAPRIMW-AVHNGTRPPLIKN 254
Db 173 MAPEVIQSLPVSETCDTYSYGVVLWEMLTREVPPFKGLEG--LQVAVLWVEKNERLTIPSS 230

QY 255 LPKPIESLMTRCWSKDPQRPSMEEIVKIMTHLMRYFPGADEPLQYPCQYSDGQNSAT 314
Db 231 CPRSFAELLHQWEADAKKRFQIISIL-----ESMSNDT- 267

QY 315 STGSPMDIANTNTSNKSDTNMEQVPATNDTIKRLESKLLKNQAKQOSESGRLSL---GAS 371
Db 268 ---SLPDKCNSFLHNKAEMRCE-IEATLERLKKLERDLSFKEQELKERERRLKMWEQKLT 323

QY 372 HGSSVESLPPTS-----EGKMSADMSEIEARIAATT-GNGQPRRRSIQDLTVTGT 421
Db 324 EQNTPLLLPLAARMSEESYFESKTESNAEMSCQITATNSGEGHGMNPSLQAMLMGF 383

QY 422 EPQVSSRSSSPSV 435
Db 384 --GDIFS MNKAGAV 395

RESULT 5
US-09-221-527-5
; Sequence 5, Application US/09221527
; Patent No. 6146832
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,527
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-527-5

Query Match      16.0%; Score 481; DB 4; Length 455;
Best Local Similarity 30.4%; Pred. No. 6.2e-29;
Matches 132; Conservative 81; Mismatches 151; Indels 70; Gaps 18;

QY 27 NFEEDYKEIEVEEVGAGFVVCCKAW--RAKDVAIK--QIESERKAFIVELROL 81
Db 7 SFVQIKFDDLQFFENC GGSGFSGVYRAKWSQDKEVAVKLLKIEKEAE-----IL 57

QY 82 SRVNHPIVKLYGACLNLP--VCLVMEYAEAGSLYNVLHG--AEPLPYVYTAHAHMSWCLOC 137
Db 58 SVLSHRNIIQFYGVILEPPNYGIVTEYASLGSLDYINSRSEEM---DMDHIMTWATDV 114

QY 138 SQGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFGTACDIQTHMTNNK--GSAAW 195
Db 115 AKGMHYLHMEAPVKVIHRDLKSRNVVIAADG-VLKICDFG-ASRPHNHTTHMSLVGTFFW 172

QY 196 MAPEVFEGSNYSEKCDVFSWGIIILWEVITRRKPFDEIGGPAPRIMW-AVHNGTRPPLIKN 254
Db 173 MAPEVIQSLPVSETCDTYSYGVVLWEMLTREVPPFKGLEG--LQVAVLWVEKNERLTIPSS 230

QY 255 LPKPIESLMTRCWSKDPQRPSMEEIVKIMTHLMRYFPGADEPLQYPCQYSDGQNSAT 314
Db 231 CPRSFAELLHQWEADAKKRFQIISIL-----ESMSNDT- 267

QY 315 STGSPMDIANTNTSNKSDTNMEQVPATNDTIKRLESKLLKNQAKQOSESGRLSL---GAS 371
Db 268 ---SLPDKCNSFLHNKAEMRCE-IEATLERLKKLERDLSFKEQELKERERRLKMWEQKLT 323

QY 372 HGSSVESLPPTS-----EGKMSADMSEIEARIAATT-GNGQPRRRSIQDLTVTGT 421
Db 324 EQNTPLLLPLAARMSEESYFESKTESNAEMSCQITATNSGEGHGMNPSLQAMLMGF 383

QY 422 EPQVSSRSSSPSV 435
Db 384 --GDIFS MNKAGAV 395

RESULT 6
US-09-221-236-5
; Sequence 5, Application US/09221236
; Patent No. 6146841
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,236
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-236-5

Query Match      16.0%; Score 481; DB 4; Length 455;
Best Local Similarity 30.4%; Pred. No. 6.2e-29;
Matches 132; Conservative 81; Mismatches 151; Indels 70; Gaps 18;

QY 27 NFEEDYKEIEVEEVGAGFVVCCKAW--RAKDVAIK--QIESERKAFIVELROL 81
Db 7 SFVQIKFDDLQFFENC GGSGFSGVYRAKWSQDKEVAVKLLKIEKEAE-----IL 57

QY 82 SRVNHPIVKLYGACLNLP--VCLVMEYAEAGSLYNVLHG--AEPLPYVYTAHAHMSWCLOC 137
Db 58 SVLSHRNIIQFYGVILEPPNYGIVTEYASLGSLDYINSRSEEM---DMDHIMTWATDV 114

QY 138 SQGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFGTACDIQTHMTNNK--GSAAW 195
Db 115 AKGMHYLHMEAPVKVIHRDLKSRNVVIAADG-VLKICDFG-ASRPHNHTTHMSLVGTFFW 172

QY 196 MAPEVFEGSNYSEKCDVFSWGIIILWEVITRRKPFDEIGGPAPRIMW-AVHNGTRPPLIKN 254
Db 173 MAPEVIQSLPVSETCDTYSYGVVLWEMLTREVPPFKGLEG--LQVAVLWVEKNERLTIPSS 230

QY 255 LPKPIESLMTRCWSKDPQRPSMEEIVKIMTHLMRYFPGADEPLQYPCQYSDGQNSAT 314
Db 231 CPRSFAELLHQWEADAKKRFQIISIL-----ESMSNDT- 267

QY 315 STGSPMDIANTNTSNKSDTNMEQVPATNDTIKRLESKLLKNQAKQOSESGRLSL---GAS 371
Db 268 ---SLPDKCNSFLHNKAEMRCE-IEATLERLKKLERDLSFKEQELKERERRLKMWEQKLT 323

QY 372 HGSSVESLPPTS-----EGKMSADMSEIEARIAATT-GNGQPRRRSIQDLTVTGT 421
Db 324 EQNTPLLLPLAARMSEESYFESKTESNAEMSCQITATNSGEGHGMNPSLQAMLMGF 383

QY 422 EPQVSSRSSSPSV 435
Db 384 --GDIFS MNKAGAV 395
```

```
Db 115 AKGMHYLHMEAPVKVIHRDLKSRNVIAADG-VLKICDFG-ASRFNHHTTMSLVGTFPW 172
QY 196 MAPEVFGSNVSEKCDVSWGIIIMEVITRRKRPDEIGCAPRIMW-AVHNKSTPPLIKN 254
Db 173 MAPEVIOQLPVSECTDYSGVGLWEMLTREVPFKGLEG--LQVAMLVVEKNERLTPSS 230
QY 255 LKPIESLMTRCWSKDPSPSMSEIYKIMTHLMRYFPGADEPLQYPCQYSDGOSNSAT 314
Db 231 CPRSFALLHQCWEADAKKRPSPFOIISIL-----ESMSNDT- 267
QY 315 STGSFMDIASTNTSKSDTNMEQVPATNDTIKRLSKLLKNQAKQOESGRSLT---GAS 371
Db 268 ---SLPDKCNFHLNKAEMWRC-IEATLERLKKLERDLSFKOELEKERERLKMWEOKLT 323
QY 372 HGSVESLPTS-----EGKMSADMSEIARIATT-GNGQPRRSIODLTVTGT 421
Db 324 EOSNTPLLPLLAARMSSESYFESKTESNSAEMSCQITATNSGEGHGNPISLOAMLMGF 383
QY 422 EPGQVSSRSSPSV 435
Db 384 --GDIFSMNKAQAV 395
```

```
RESULT 7
US-09-221-416-5
; Sequence 5, Application US/09221416
```

```
; Patent No. 6153417
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/221,416
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-416-5
```

```
Query Match 16.0%; Score 481; DB 4; Length 455;
Best Local Similarity 30.4%; Pred. No. 6,2e-29;
Matches 132; Conservative 81; Mismatches 151; Indels 70; Gaps 18;
```

```
QY 27 NFEIDYKEIEVEEVVGRGAFVVCYAKM--RAKDVAIK---QIESESRKAFIVELRQL 81
Db 7 SFVQIKFDDLOFPENCGGSGFSGSVYRAKWTISQDEKVAVKKLRIEKEAE-----IL 57
QY 82 SRVNHPIVLYGACLN--VCLVMEYAEGSLYNYLHG--AEPLPYTAAHAMSWCLOC 137
Db 58 SVLSHRNITQFYVILPEPNYGIATEYASLSGLYDIYNSRSEEM--DMDHMTWATDV 114
QY 138 SOGVAAYLHSMQPKALIHRLDKPPLLIVAGTVLKIICDFGTACDIQTHMTNKK--GSAAM 195
Db 115 AKGMHYLHMEAPVKVIHRDLKSRNVIAADG-VLKICDFG-ASRFNHHTTMSLVGTFPW 172
QY 196 MAPEVFGSNVSEKCDVFSNGIILMEVITRRKRPDEIGCAPRIMW-AVHNKSTPPLIKN 254
Db 173 MAPEVIOQLPVSECTDYSGVGLWEMLTREVPFKGLEG--LQVAMLVVEKNERLTPSS 230
QY 255 LKPIESLMTRCWSKDPSPSMSEIYKIMTHLMRYFPGADEPLQYPCQYSDGOSNSAT 314
Db 231 CPRSFALLHQCWEADAKKRPSPFOIISIL-----ESMSNDT- 267
QY 315 STGSFMDIASTNTSKSDTNMEQVPATNDTIKRLSKLLKNQAKQOESGRSLT---GAS 371
Db 268 ---SLPDKCNFHLNKAEMWRC-IEATLERLKKLERDLSFKOELEKERERLKMWEOKLT 323
QY 372 HGSVESLPTS-----EGKMSADMSEIARIATT-GNGQPRRSIODLTVTGT 421
```

```
Db 324 EOSNTPLLPLLAARMSSESYFESKTESNSAEMSCQITATNSGEGHGNPISLOAMLMGF 383
QY 422 EPGQVSSRSSPSV 435
Db 384 --GDIFSMNKAQAV 395
```

```
RESULT 8
US-09-221-245-5
; Sequence 5, Application US/09221245
```

```
; Patent No. 6180358
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/221,245
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: US 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-245-5
```

```
Query Match 16.0%; Score 481; DB 4; Length 455;
Best Local Similarity 30.4%; Pred. No. 6,2e-29;
Matches 132; Conservative 81; Mismatches 151; Indels 70; Gaps 18;
```

```
QY 27 NFEIDYKEIEVEEVVGRGAFVVCYAKM--RAKDVAIK---QIESESRKAFIVELRQL 81
Db 7 SFVQIKFDDLOFPENCGGSGFSGSVYRAKWTISQDEKVAVKKLRIEKEAE-----IL 57
QY 82 SRVNHPIVLYGACLN--VCLVMEYAEGSLYNYLHG--AEPLPYTAAHAMSWCLOC 137
Db 58 SVLSHRNITQFYVILPEPNYGIATEYASLSGLYDIYNSRSEEM--DMDHMTWATDV 114
QY 138 SOGVAAYLHSMQPKALIHRLDKPPLLIVAGTVLKIICDFGTACDIQTHMTNKK--GSAAM 195
Db 115 AKGMHYLHMEAPVKVIHRDLKSRNVIAADG-VLKICDFG-ASRFNHHTTMSLVGTFPW 172
QY 196 MAPEVFGSNVSEKCDVFSNGIILMEVITRRKRPDEIGCAPRIMW-AVHNKSTPPLIKN 254
Db 173 MAPEVIOQLPVSECTDYSGVGLWEMLTREVPFKGLEG--LQVAMLVVEKNERLTPSS 230
QY 255 LKPIESLMTRCWSKDPSPSMSEIYKIMTHLMRYFPGADEPLQYPCQYSDGOSNSAT 314
Db 231 CPRSFALLHQCWEADAKKRPSPFOIISIL-----ESMSNDT- 267
QY 315 STGSFMDIASTNTSKSDTNMEQVPATNDTIKRLSKLLKNQAKQOESGRSLT---GAS 371
Db 268 ---SLPDKCNFHLNKAEMWRC-IEATLERLKKLERDLSFKOELEKERERLKMWEOKLT 323
QY 372 HGSVESLPTS-----EGKMSADMSEIARIATT-GNGQPRRSIODLTVTGT 421
Db 324 EOSNTPLLPLLAARMSSESYFESKTESNSAEMSCQITATNSGEGHGNPISLOAMLMGF 383
QY 422 EPGQVSSRSSPSV 435
Db 384 --GDIFSMNKAQAV 395
```

```
RESULT 9
US-09-163-115-5
; Sequence 5, Application US/09163115A
; Patent No. 6183962
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
```

```

; CURRENT APPLICATION NUMBER: US/09/163,115A
; CURRENT FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-163-115-5

```

Query Match	16.0%;	Score 481;	DB 4;	Length 455;
Best Local Similarity	30.4%;	Pred. NO. 6.2e-29;		
Matches 132;	Conservative	81;	Mismatches 151;	Indels 70; Gaps 18;

QY	27	NFEEDYKIEVEEVUGAPGVCKAKW---RAKVAIK---QIESERKAFIVELRQL	81
Db	7	SRVQIKFDPOFFENCGGSGFSYRAKWTISQKEVAVKLLKIEKEAE-----IL	57
QY	82	SRVNHPIVTKLYGACLANP--VCLVMEYAEAGSLYNNVLHG--AEPLPYTYAAHAMSWCLOC	137
Db	58	SVLSHRNLIQFYGVILPFPNYGIVTEVASLGSUYDIYNSRSEEN---DMDHIMTWATDV	114
QY	138	SGOVAYLHSMOPKALIHRLDKPNNLLVAGGTGVLKICDFTGACDIQTHTMNNK--GSAAW	195
Db	115	AKGMHYLHWEAPVKVIHRLDKSRNVVIAADG-VLKICDFG-ASRPHNHTTHWSLVGTFPW	172
QY	196	MAPEVFEGSNSYSEKCDVFSWGIIILWEVITRPFDEIGGPAFRIMW-AVHNGTRPPLIKN	254
Db	173	MAPEVIQSLPVSFETCDTYSYGVVLWEMLTREVFPFKGLEG--LQVAWLVEKNERLTI	230
QY	255	LPKPTIESLMTRCWSKDPQRSQPSMEEIVKIMTHLMRYFPGADEPLQYPCQYDEGQNSAT	314
Db	231	CPRSFAELLHQWEADAKRPSFKQIISIL-----ESMSNDT-267	
QY	315	STGSPMDIASTNTNKGSDTNWEQVPATNDTIKRLSKLLKNOAKQOSESRLSL---GAS	371
Db	268	---SUPDKCNSFLHNKAWRCE-IEATLERLKLKLERDLFSKEQELKERERLKKWEOKLT	323
QY	372	HGSSVESLPPTS-----EGKRSADMSIEIARIAATT-GNCQPRRSIODLTVGT	421
Db	324	EOSNTPLLLPLAARMSSESFESKTESNAEMSQITATSNGEHGHNPSLOQNMJMGF	383
QY	422	EPGQVSSRSSSPSV	435
Db	384	--GDIFSNNKAGAV	395

```

RESULT 10
US-09-221-528-5
; Sequence 5, Application US/09221528
; Patent No. 6190874
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,528
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-528-5

```

Query Match 16.0%; Score 481; DB 4; Length 455;
Best Local Similarity 30.4%; Pred. No. 6.2e-29;
Matches 132; Conservative 81; Mismatches 151; Indels 70; Gaps 18;

QY 27 NFEEIDYKEIEVEEVVGRGAFVCKAKW--RAKDVAIK--QIESESEKAFIVELROL 81

Db	7	SFVQIKFDLLOFFENCGGSGFSVYRAKWI	SQDKEVAVKLLKIEKAE-----	IL	57
Qy	82	SRVNHPNIVKLYGACINP--VCLWMEV	AEAGSLYNVLHG--ABEPLPYTTAAHMSWCLOC	137	
Db	58	SVLSHRNIIQFYGVILEPPNNGIVTEYAS	LSGLSYDYINSNRSEEM---DMDHIMTWATDV	114	
Qy	138	SOQVAYLHSMOPKALIHRLDKPNNLLVAG	STVLKICDFTACDIQTHMTNNK--GSAAW	195	
Db	115	AKGWHYLHMEAPVKVIRHRLDKSRNVIAA	DG-VLKICDFF-ASRFHNHTTHMSLVGTFFP	172	
Qy	196	MAPEVFGSNYSEKCDVFSWGIIILWEVIT	RRKPFDEIGGPAFRIMW-AVHNGTRPPLIKN	254	
Db	173	MAPEVIOQLPVSETCDYISGVVLWELT	TEVPFKLEG--LOVAMLVVEKNERLTIPSS	230	
Qy	255	LPKPIESIMTRCWSKDPSPQRSPEEIVKI	IMTHLMRYPPGADEPLQPCQYSDGQGSNAT	314	
Db	231	CPRSFALLHCQWEADAKKPPSPFOIISIL	-----ESMNGDT-	267	
Qy	315	STGSFMDIASTNTSNKSDTNMEOVPATNDT	IKRLSEKLLKNQAKQOSESGRLSL---GAS	371	
Db	268	--SLPDKCNSFLHNKAERCE-IEATLERL	KKLRLDLSPKEQELKERERLKNWEQKLT	323	
Qy	372	HGSSVESLPPTS-----EKRMSADMSEI	EARIAATT-NGQFRRRSIODLTVTGT	421	
Db	324	EQSNTPULLLAARMSEESYFESKTEBSNA	ESCOIATSNGECHGMNPSLOAMLMGF	383	
Qy	422	EPGQVSSRSSSPSV	435		
Db	384	--GDIFSMNKAGAV	395		

RESULT 17

```

US-09-593-553-5
; Sequence 5, Application US/09593553
; Patent No. 6200770
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09-
; CURRENT FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: 09/163,
; PRIOR FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-593-553-5

```

Query Match 16.0%; Score 481; DB 4; Length 455;
Best Local Similarity 30.4%; Pred. NO. 6.2e-29;
Matches 132; Conservative 81; Mismatches 151; Indels 70; Gaps 18;

Qy	27	NFBEIDYKETEVEEVGRGAFGVVCKAKW--RAKDVAIK---QIESERKAFIVELROL	81
Db	7	SFVQIKFDLQFENCNGGSGFGSVYKAKISQDKEVAVKLLKIEKEAE-----IL	57
Qy	82	SRVNHPIVVKLYGACILNP--VCLVMYEYAGSGSLYNNVLHG--AEPILPYITAAHAMSMCLQC	137
Db	58	SVLSHRNIIQFYGVILEPPNYGIVTEYASLSGLSDYVINSNRSEEM---DMDHIMTIAWDV	114
Qy	138	SOGVAYLHSMQPKALHTRDLKPNELLLVAGGTVLKLCDFEGTACDIOTHTNNK--GSAAW	195
Db	115	AKGMHYLHMEAPVKVILHTRDLKSRNVVIADG-VLKLCDFG-ASRFHNHTTHMSLVGTFFPW	172
Qy	196	MAPEVEGSGNSYSEKCDVFWGIIILMEVITTRKPFDEIGGPAFRMW-AVHNGTRPLIKN	254
Db	173	MAPEVIQSUPVSTCTDYGVVVLEMLTRVFPFKELEG--EQVAVLWVVEKNERUTIPSS	230
Qy	255	LPKPIESLTRCWKSDPORSORMEETIVKIMTLHMYFPGADEPLQVPCOYSDGEGQSNAT	314

Db 231 CRRSEAEHLHQCEWADAKRPSFKQIISIL-----ESMSNDT- 267
QY 315 STGSEMDIASTNTSKSDTNEQVPATNDTIKRLSKLLKNQAKQSESGRLSL---GAS 371
Db 268 ---SLPDKCNLSLHNAEMRCE-IEATLERLKKLERDLSFEGQELKEERLKKMEQKLT 323
QY 372 HGSSEVESLPPTS-----EGKMSADMSIEIARIAATT-GNGQPRRSIOTLTVTGT 421
Db 324 EGSNTPULLPLAARMSEESYFESKTEESNSAEMSCQITATNSGEGHGMNPISLOAMLMG 383
QY 422 EPGQVSSRSSSPSV 435
Db 384 --GDIFSMNKAGAV 395

RESULT 12
US-09-221-237-5
; Sequence 5, Application US/09221237
; Patent No. 6214597
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNT-050
; CURRENT APPLICATION NUMBER: US/09/221,237
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-221-237-5

Query Match 16.0%; Score 481; DB 4; Length 455;

Best Local Similarity 30.4%; Pred. No. 6, 2e-29;
Matches 132; Conservative 81; Mismatches 151; Indels 70; Gaps 18;

QY 27 NFEIIEYKEIEVEEVGAGFVVCCKAKM--RAKQVAIK--QIESESRKAFIVELROL 81
Db 7 SFOQIKFDDLDQFPENCNGGSGFSVYRAKMI SQKEVAVAKKLIKIEKAL-----IL 57
QY 82 SRVNHFNIVLYGACLN--VCLVMEYABGSLYNVLAG--AEPLPYTAHAMSWCLOQC 137
Db 58 SVLSHNIIQFYGVILLEPPRYGIVTEYASLSGLDYINSRSEB---DMDHIMTATDV 114
QY 138 SOGVAIYHSMQKALIHRLDKPRLNLLVAGTYLKTCDGTACDICTHMTNKK--GSAAM 195
Db 115 AKGMHYLHMEAPVKVYIHRDLKSRNVVIAADG-VLKI CD FG-ASRFHHTHMSLVGTFFP 172
QY 196 MAPEVEGSSVSKCDVFSWGIIIMEVITRRKPFDEIGGPAFRIM-AVHNGTRPPLIKN 254
Db 173 MAPEVIOSLPVSCTDTYISGVVLMEMLTREVPKGLBEG--LOVALLVKEKNERLTPSS 230
QY 255 LKPRISLITRCSKDPQSRPSMEIYKIMTHLMRYFPGADEPLQYPCQYSDGGSNATF 314
Db 231 CRRSEAEHLHQCEWADAKRPSFKQIISIL-----ESMSNDT- 267
QY 315 STGSEMDIASTNTSKSDTNEQVPATNDTIKRLSKLLKNQAKQSESGRLSL---GAS 371
Db 268 ---SLPDKCNLSLHNAEMRCE-IEATLERLKKLERDLSFEGQELKEERLKKMEQKLT 323
QY 372 HGSSEVESLPPTS-----EGKMSADMSIEIARIAATT-GNGQPRRSIOTLTVTGT 421
Db 324 EGSNTPULLPLAARMSEESYFESKTEESNSAEMSCQITATNSGEGHGMNPISLOAMLMG 383
QY 422 EPGQVSSRSSSPSV 435
Db 384 --GDIFSMNKAGAV 395

RESULT 13

US-09-291-839-2
; Sequence 2, Application US/09291839A
; Patent No. 6261818
; GENERAL INFORMATION:
; APPLICANT: Raju, Jayaseelan
; TITLE OF INVENTION: NOVEL CARC PROTEIN AND NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNT-068
; CURRENT APPLICATION NUMBER: US/09/291,839A
; CURRENT FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 835
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-291-839-2

Query Match 13.9%; Score 418; DB 4; Length 835;
Best Local Similarity 31.6%; Pred. No. 9e-24;
Matches 115; Conservative 67; Mismatches 144; Indels 38; Gaps 13;

QY 31 IDYKEIEVEEVGAGFVVCCKAKMRAKQVAIKQIE-----SESESRKAFIVELROL 85
Db 458 LQISEIEFHILIGSGFGKXKRCNKIYAIKRYRANTYCSKSDVDFCREVSLICQLN 517
QY 86 HPNIVLYGACLN--PVCVMEYABGSLYNVLAGAEPLPYTAHAMSWCLOCSQVA 142
Db 518 HPCVIOFVGACLNPPQFALVTQYISGSLFSLH--EQKRILDIQSKLIIVAVAKME 575
QY 143 YHSM-OPKALIHRLDKPRLNLLVAGTYLKTCDGTACDICT---HMTNKGSAAMNA 197
Db 576 YHNLTOP--IHRDLNSHILLYEDGHV-VADFGESRLOSLDDNMTKQCGNLRMA 632
QY 198 PEVF-EGSNYSEKCDVFSWGIIIMEVITRRKPFDEIGGPAFRIMAVHNGTRPPLIKNL 256
Db 633 PEVFTQCTRYTIKADVFSYALCLWEILITGEIPRAHLKPAADAMVYH-IRPPIGYSIP 691
QY 257 KPISLITRCSKDPQSRPSMEIYKIMTHLMRYFPGADPEPLQYPCQYSDGGSNATF 316
Db 692 KPISLITRCSKDPQSRPSMEIYKIMTHLMRYFPGADPEPLQYPCQYSDGGSNATF 316
QY 317 GSFMDIASTN-TSKSDTNEQVPATNDTIKRLSKLLKNQAKQSESGRLSIGASHGSS 375
Db 738 GSLSPSSSDCLVNRGPGSRSHVALLSRPELEVALLNARSYALLSQSAGYS---SGLS 794
QY 376 VESL 379
Db 795 LEM 798

RESULT 14
US-08-205-018-2
; Sequence 2, Application US/08205018
; Patent No. 5554523
; GENERAL INFORMATION:
; APPLICANT: Reddy, Usharani R.
; TITLE OF INVENTION: No. 5554523e1 Protein Kinase, Nucleic Acid
; TITLE OF INVENTION: Sequences Encoding the Same and Methods Related
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and
; ADDRESS: No. 5554523e1s
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/205,018
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Gaumond, Rebecca R.
; REGISTRATION NUMBER: 35,152
; REFERENCE/DOCKET NUMBER: CH-0488
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 668 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-205-018-2

Query Match 13.7%; Score 412; DB 1; Length 668;
Best Local Similarity 24.6%; Pred. No. 1.9e-23;
Matches 133; Conservative 79; Mismatches 181; Indels 148; Gaps 20;

QY 30 EIDYKEIEVEVGAGFVGVCKAKWRAKDVAIKQIESESEKAFIVELRQLSRVNHPI 89
DB 119 EVPPEILDQLQWVGSGAQQAVFLGRFHGEEVAVKVRDLKE-----TDIKHLRKLKHPNI 173
QY 90 VKLYGACINP--VCLVMEYAEGLSYNLVHGAEPLPYTTAAHAMSWCLQCSQGVAYLHSM 147
DB 174 ITFKGVCTQAPCYCILMEFCAQQQLYEVLRAGRPV---TPSLLDVWSMGTAGMNYLHLH 230
QY 148 QPKALIHRDLKPPNLLVAGTGLVKICDFTACDIQTHMTNNK--GSAAMWAEVFEESN 205
DB 231 K---IIHRDLKSPN-MLITYDDVVKISDFGTSKELSDKSTKMSFAGTVAMWAEVIRNEP 286
QY 206 YSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMWAV-HNGTRPPLIKNLPKPIESLMT 264
DB 287 VSEKVDIWSFGVWLWELLTGEIPYKVDSSA--IIVGVGNSLHLPVPSSCPDGFKILLR 344
QY 265 RCWSKDPQRSMBEIVKIMTHLMRYFPGADEPLOYPCQYSDGQSNSTATSGSFMDIAS 324
DB 345 QCWNKSPNRPFRQ---ILLHL-----DIAS 368
QY 325 TNT-SNKSDT-----NMQVPATNDTIKRLSKLL----- 353
DB 369 ADVLSTPQETVFKSQAEWREVKLHFKEIKSEGTCLHRLHEELVMRRRELRHALDIREH 428
QY 354 -----PPTSEGRMSADMEIEARI-----AATTGQPRRRSIQDLTVGTGEPQVSS 428
DB 429 YERKLERANNLYMELNALMLQLELKERELLRREQALERRCFGLLKPHPSRGLLHGNTMEK 488
QY 379 L-----PPTSEGRMSADMEIEARI-----AATTGQPRRRSIQDLTVGTGEPQVSS 428
DB 489 LIKKRNPQNLSPHSRQPDILKAESLLPKLDAALSGVGLP-----GCPKAPPSGR--S 540
QY 429 RSSSPSVRMITTSPTSEKP---TRSHPTPDDSDTDNG-----SDNSIPMAVLTLDHQ 479
; TITLE OF INVENTION: Sequences Encoding the Same and Methods Related Thereto
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESS: Woodcock Washburn Kurtz Mackiewicz and No. 5676945x1s
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch disk, 720 Kb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/395,580
; FILING DATE: herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/205,018
; FILING DATE: 01-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Rebecca L. Ralph (formerly Gaumond)
; REGISTRATION NUMBER: 35,152
; REFERENCE/DOCKET NUMBER: CH-0488
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 859 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-395-580-2

Query Match 13.7%; Score 412; DB 1; Length 859;
Best Local Similarity 24.6%; Pred. No. 2.7e-23;
Matches 133; Conservative 79; Mismatches 181; Indels 148; Gaps 20;

QY 30 EIDYKEIEVEVGAGFVGVCKAKWRAKDVAIKQIESESEKAFIVELRQLSRVNHPI 89
DB 119 EVPPEILDQLQWVGSGAQQAVFLGRFHGEEVAVKVRDLKE-----TDIKHLRKLKHPNI 173
QY 90 VKLYGACINP--VCLVMEYAEGLSYNLVHGAEPLPYTTAAHAMSWCLQCSQGVAYLHSM 147
DB 174 ITFKGVCTQAPCYCILMEFCAQQQLYEVLRAGRPV---TPSLLDVWSMGTAGMNYLHLH 230
QY 148 QPKALIHRDLKPPNLLVAGTGLVKICDFTACDIQTHMTNNK--GSAAMWAEVFEESN 205
DB 231 K---IIHRDLKSPN-MLITYDDVVKISDFGTSKELSDKSTKMSFAGTVAMWAEVIRNEP 286
QY 206 YSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMWAV-HNGTRPPLIKNLPKPIESLMT 264
DB 287 VSEKVDIWSFGVWLWELLTGEIPYKVDSSA--IIVGVGNSLHLPVPSSCPDGFKILLR 344
QY 265 RCWSKDPQRSMBEIVKIMTHLMRYFPGADEPLOYPCQYSDGQSNSTATSGSFMDIAS 324
DB 345 QCWNKSPNRPFRQ---ILLHL-----DIAS 368
QY 325 TNT-SNKSDT-----NMQVPATNDTIKRLSKLL----- 353
DB 369 ADVLSTPQETVFKSQAEWREVKLHFKEIKSEGTCLHRLHEELVMRRRELRHALDIREH 428
QY 354 -----PPTSEGRMSADMEIEARI-----AATTGQPRRRSIQDLTVGTGEPQVSS 428
DB 429 YERKLERANNLYMELNALMLQLELKERELLRREQALERRCFGLLKPHPSRGLLHGNTMEK 488
QY 379 L-----PPTSEGRMSADMEIEARI-----AATTGQPRRRSIQDLTVGTGEPQVSS 428
DB 489 LIKKRNPQNLSPHSRQPDILKAESLLPKLDAALSGVGLP-----GCPKAPPSGR--S 540
QY 429 RSSSPSVRMITTSPTSEKP---TRSHPTPDDSDTDNG-----SDNSIPMAVLTLDHQ 479
; GENERAL INFORMATION:
; APPLICANT: Usharani R. Reddy, David Pleasure and the Children's
; APPLICANT: Hospital of Philadelphia
; TITLE OF INVENTION: No. 5676945el Protein Kinase, Nucleic Acid
```

Db 541 RRGKTRHRKASAKSCGDLPGJLRTAVPHEPGSPGSPGJGSGPSAMEACPPALRGJLHD 600
Qy 480 L 480
Db 601 L 601

Search completed: December 9, 2002, 23:00:36
Job time : 28.5928 secs

GenCore version 5.1.1.3
 Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 9, 2002, 22:53:24 ; Search time 116.014 Seconds
 (without alignments)
 81.062 Million cell updates/sec

Title: US-09-830-144-2
 Perfect score: 3014
 Sequence: 1 MSTAASASSSSSSAGEMIE.....CKKQLEIVRSQQKRGQTS 579

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 103943 seqs, 16242309 residues

Total number of hits satisfying chosen parameters: 103943

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pep.*
 2: /cgn2_6/prodata/2/pubpaa/ECT_NEW_PUB.pep.*
 3: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pep.*
 4: /cgn2_6/prodata/2/pubpaa/US06_PUBCOMB.pep.*
 5: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep.*
 6: /cgn2_6/prodata/2/pubpaa/US07_PUBCOMB.pep.*
 7: /cgn2_6/prodata/2/pubpaa/ECTUS_PUBCOMB.pep.*
 8: /cgn2_6/prodata/2/pubpaa/US08_PUBCOMB.pep.*
 9: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pep.*
 10: /cgn2_6/prodata/2/pubpaa/US03_PUBCOMB.pep.*
 11: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pep.*
 12: /cgn2_6/prodata/2/pubpaa/US10_PUBCOMB.pep.*
 13: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pep.*
 14: /cgn2_6/prodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3014	100.0	579	US-10-158-895-4	Sequence 4, Appli
2	3014	100.0	590	US-10-158-895-15	Sequence 15, Appli
3	481	16.0	455	US-09-757-982-5	Sequence 5, Appli
4	461.5	15.3	1036	US-10-014-882-2	Sequence 2, Appli
5	418.5	13.9	394	US-09-862-027-19	Sequence 19, Appli
6	418	13.9	835	US-09-947-199-2	Sequence 2, Appli
7	416	13.8	328	US-03-862-027-18	Sequence 18, Appli
8	414	13.7	835	US-09-947-199-8	Sequence 8, Appli
9	408.5	13.6	966	US-09-771-161A-197	Sequence 197, App
10	393	13.0	263	US-09-840-704-5	Sequence 5, Appli
11	392	13.0	850	US-09-904-389-2	Sequence 2, Appli
12	369	12.2	425	US-03-828-313-29	Sequence 29, Appli
13	363.5	12.1	527	US-09-977-269-10	Sequence 10, Appli
14	360.5	12.0	277	US-09-882-166-4	Sequence 4, Appli
15	359	11.9	278	US-09-842-582-4	Sequence 13, Appli
16	359	11.9	278	US-09-797-039-13	Sequence 18, Appli
17	359	11.9	278	US-09-922-138-18	Sequence 27, Appli
18	359	11.9	278	US-09-922-138-27	Sequence 17, Appli
19	359	11.9	278	US-09-910-150-17	

ALIGNMENTS

RESULT 1

US-10-158-895-4
 ; Sequence 4, Application US/10158895
 ; Patent No. US20020155624A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ONO, KOICHIRO
 ; APPLICANT: TSUCHIYA, MASAYUKI
 ; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
 ; FILE REFERENCE: 053466/0278
 ; CURRENT APPLICATION NUMBER: US/10/158,895
 ; CURRENT FILING DATE: 2002-06-03
 ; PRIOR APPLICATION NUMBER: US/09/529,279
 ; PRIOR FILING DATE: 2000-04-11
 ; PRIOR APPLICATION NUMBER: PCT/JP98/04796
 ; PRIOR FILING DATE: 1998-10-22
 ; PRIOR APPLICATION NUMBER: JP 9/290188
 ; PRIOR FILING DATE: 1997-10-22
 ; NUMBER OF SEQ ID NOS: 48
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 4
 ; LENGTH: 579
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-158-895-4

Query Match 100.0%; Score 3014; DB 9; Length 579;
 Best Local Similarity 100.0%; Pred. No. 8,1e-176;
 Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTAASASSSSSSAGMIAPSOVLNFEIDYKEIEVEEVGRGAFGVVCKAKRAKDV 60
 |||||
 DB 1 MSTAASASSSSSSAGMIAPSOVLNFEIDYKEIEVEEVGRGAFGVVCKAKRAKDV 60
 |||||

QY 61 AIKQIESESRKAFIVELRQLSRVNHPIVKLYGACLNVPCLVMEYAEAGSLYNVLHGA 120
 |||||
 DB 61 AIKQIESESRKAFIVELRQLSRVNHPIVKLYGACLNVPCLVMEYAEAGSLYNVLHGA 120
 |||||

QY 121 PLPYTTAAHMSWCLOCSQGVAYLHSMQPKALIHRLDKPNNLLLVAGGTVLKICDFGTAC 180
 |||||
 DB 121 PLPYTTAAHMSWCLOCSQGVAYLHSMQPKALIHRLDKPNNLLLVAGGTVLKICDFGTAC 180
 |||||

QY 181 DIQTHMTNKGSAAMWAPEVEFGSNYSEKCDVFSWGIILMEVITRRKPFDEIGPAFRIM 240
DB 181 DIQTHMTNKGSAAMWAPEVEFGSNYSEKCDVFSWGIILMEVITRRKPFDEIGPAFRIM 240
QY 241 WAVHNGTRPPLIKNLPKPIESIMTRCWSKDPQSPRSMIEIVKIMTHLMRYFGADEPLQY 300
DB 241 WAVHNGTRPPLIKNLPKPIESIMTRCWSKDPQSPRSMIEIVKIMTHLMRYFGADEPLQY 300
QY 301 PCQYSDGQSNASTSGSFMDIASTNTSNKSDPTNMEQVATNDTIKRLSKLLKNQAKQ 360
DB 301 PCQYSDGQSNASTSGSFMDIASTNTSNKSDPTNMEQVATNDTIKRLSKLLKNQAKQ 360
QY 361 SESGRSLGASHGSSVESLPPTSEGRKMSADMEIARIATGNGQPRRSIIDLTVTG 420
DB 361 SESGRSLGASHGSSVESLPPTSEGRKMSADMEIARIATGNGQPRRSIIDLTVTG 420
QY 421 TEPQVSSRSSSPSVMTTSGPTSEKPTRSHPWTPDDSDTDNGSDNSIPMAVLTLDHQL 480
DB 421 TEPQVSSRSSSPSVMTTSGPTSEKPTRSHPWTPDDSDTDNGSDNSIPMAVLTLDHQL 480
QY 481 QPLAPCPNKSMAVFEQHCMAQOEYMKVQTEIALLQKQELVABLDQDEKQDQNTSRL 540
DB 481 QPLAPCPNKSMAVFEQHCMAQOEYMKVQTEIALLQKQELVABLDQDEKQDQNTSRL 540
QY 541 VOEHKKULDENKSLSTYYQCKQLEVIRSQQKRGTS 579
DB 541 VOEHKKULDENKSLSTYYQCKQLEVIRSQQKRGTS 579

RESULT 2
US-10-158-895-15
Sequence 15, Application US/10158895
Patent No. US20020155624A1
GENERAL INFORMATION:
APPLICANT: ONO, KOICHIRO
APPLICANT: OHTOMO, TOSHIHIKO
APPLICANT: TSUCHIYA, MASAYUKI
TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
FILE REFERENCE: 053466/0278
CURRENT APPLICATION NUMBER: US/10/158, 895
PRIOR FILING DATE: 2002-06-03
PRIOR APPLICATION NUMBER: US/09/529, 279
PRIOR FILING DATE: 2000-04-11
PRIOR APPLICATION NUMBER: PCT/JP98/04796
PRIOR FILING DATE: 1998-10-22
PRIOR APPLICATION NUMBER: JP 9/290188
PRIOR FILING DATE: 1997-10-22
NUMBER OF SEQ ID NOS: 48
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 15
LENGTH: 590
TYPE: PRT
ORGANISM: Homo sapiens
US-10-158-895-15

Query Match 100.0%; Score 3014; DB 9; Length 590;
Best Local Similarity 100.0%; Pred. No. 8.3e-176;
Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTASASSSSSSAGMTIAPSOVUNFEIIDEKEIEVEEVGRGAFGVYCKAKMRAKDY 60
DB 1 MSTASASSSSSSAGMTIAPSOVUNFEIIDEKEIEVEEVGRGAFGVYCKAKMRAKDY 60
QY 61 AIKQIESESERKAFIYELRQLSRNHPNIVKLYGACLNPCVLMVEYAEAGSLYNVLHGAE 120
DB 61 AIKQIESESERKAFIYELRQLSRNHPNIVKLYGACLNPCVLMVEYAEAGSLYNVLHGAE 120
QY 121 PLPYTTAAHMSWCLQCSQGVAYIHSWQPKALIHRLDKPNNLLVAGGYLTKICDGTAC 180
DB 121 PLPYTTAAHMSWCLQCSQGVAYIHSWQPKALIHRLDKPNNLLVAGGYLTKICDGTAC 180
QY 181 DIQTHMTNKGSAAMWAPEVEFGSNYSEKCDVFSWGIILMEVITRRKPFDEIGPAFRIM 240

DB 181 DIQTHMTNKGSAAMWAPEVEFGSNYSEKCDVFSWGIILMEVITRRKPFDEIGPAFRIM 240
QY 241 WAVHNGTRPPLIKNLPKPIESIMTRCWSKDPQSPRSMIEIVKIMTHLMRYFGADEPLQY 300
DB 241 WAVHNGTRPPLIKNLPKPIESIMTRCWSKDPQSPRSMIEIVKIMTHLMRYFGADEPLQY 300
QY 301 PCQYSDGQSNASTSGSFMDIASTNTSNKSDPTNMEQVATNDTIKRLSKLLKNQAKQ 360
DB 301 PCQYSDGQSNASTSGSFMDIASTNTSNKSDPTNMEQVATNDTIKRLSKLLKNQAKQ 360
QY 361 SESGRSLGASHGSSVESLPPTSEGRKMSADMEIARIATGNGQPRRSIIDLTVTG 420
DB 361 SESGRSLGASHGSSVESLPPTSEGRKMSADMEIARIATGNGQPRRSIIDLTVTG 420
QY 421 TEPQVSSRSSSPSVMTTSGPTSEKPTRSHPWTPDDSDTDNGSDNSIPMAVLTLDHQL 480
DB 421 TEPQVSSRSSSPSVMTTSGPTSEKPTRSHPWTPDDSDTDNGSDNSIPMAVLTLDHQL 480
QY 481 QPLAPCPNKSMAVFEQHCMAQOEYMKVQTEIALLQKQELVABLDQDEKQDQNTSRL 540
DB 481 QPLAPCPNKSMAVFEQHCMAQOEYMKVQTEIALLQKQELVABLDQDEKQDQNTSRL 540
QY 541 VOEHKKULDENKSLSTYYQCKQLEVIRSQQKRGTS 579
DB 541 VOEHKKULDENKSLSTYYQCKQLEVIRSQQKRGTS 579

RESULT 3
US-09-757-982-5
Sequence 5, Application US/09757982
Patent No. US20020094559A1
GENERAL INFORMATION:
APPLICANT: Acton, Susan
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
FILE REFERENCE: NMI-050
CURRENT APPLICATION NUMBER: US/09/757, 982
PRIOR FILING DATE: 2001-01-10
PRIOR APPLICATION NUMBER: 09/163, 115
PRIOR FILING DATE: 1998-09-29
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 455
TYPE: PRT
ORGANISM: Homo sapiens
US-09-757-982-5

Query Match 16.0%; Score 481; DB 10; Length 455;
Best Local Similarity 30.4%; Pred. No. 2.3e-22;
Matches 132; Conservative 81; Mismatches 151; Indels 70; Gaps 18;

QY 27 NFEIIDEKEIEVEEVGRGAFGVYCKAKM--RAKDAIK--QIESESERKAFIYELRQ 81
DB 7 SFVQIKFDDLOFFENCGGSGFYRAKWTISODEKAVAKLKLKEKAE-----IL 57
QY 82 SRVHNPINVKLYGACLNPCVLMVEYAEAGSLYNVLHG--ABPLPYTTAAHMSWCLQ 137
DB 58 SVLSHRNLTQYGYILEBPNYGVITEYASLQSLYDYNRSRSEM--DMDHMTATDV 114
QY 138 SQGVAYIHSWQPKALIHRLDKPNNLLVAGGYLTKICDGTACDIQTHMTNKK--GSAAW 195
DB 115 AKGMHYLMEMAPVAVIHRDLKSRNVVIAADG-VLKIDFG-ASRFHHTTHMSLVGTFPW 172
QY 196 MAPEVEFGSNYSEKCDVFSWGIILMEVITRRKPFDEIGPAFRIM-AVHNGTRPPLIKN 254
DB 173 MAPEVIGSLPYSEICDITYSYGVVLMEMLTREVFPKLEG--LQVAMLVKERNELTIPSS 230
QY 255 LKPIESIMTRCWSKDPQSPRSMIEIVKIMTHLMRYFGADEPLQYPCQYSDGQSNAST 314
DB 231 CFSRAELHQCWEADAKRPSFKQIISIL-----ESMSDNT- 267
QY 315 STGSFMDIASTNTSNKSDPTNMEQVATNDTIKRLSKLLKNQAKQSDGRLSL--GAS 371


```
Db 268 ---SLPKCNFLHNKAERCE-IEATLERLKKLERDLSPKEQLKERERRRLLKMMEQKLT 323
QY 372 HGSVESLPPTS-----EGKMSADMSEIEARIAATT-GNGQPRRRSIQDLTVTGT 421
Db 324 EQSNTPLLLPLAARMSEESYFESKTEESNAEMSCQITATSNGEGHGMNPSLOAMWLMGF 383
QY 422 EPQGVSRSSSPSV 435
Db 384 --GDIFSMNKAGAV 395

RESULT 4
US-10-014-882-2
; Sequence 2, Application US/10014882
; Patent No. US20020107384A1
; GENERAL INFORMATION:
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James
; APPLICANT: Donoho, Gregory
; TITLE OF INVENTION: No. US20020107384A1el Human Kinase and Polynucleotides Encoding t
; FILE REFERENCE: LEX-0279-USA
; CURRENT APPLICATION NUMBER: US/10/014,882
; CURRENT FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: US 60/254,744
; PRIOR FILING DATE: 2000-12-11
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1036
; TYPE: PRT
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1036)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-014-882-2

Query Match 15.3%; Score 461.5; DB 12; Length 1036;
Best Local Similarity 26.7%; Pred. No. 9e-21;
Matches 171; Conservative 97; Mismatches 254; Indels 119; Gaps 26;

QY 22 PSVLNFEIDYKEIEVEVVGCGAFVVCCKAKRAKVAIKQIESERKAFIV----- 76
Db 110 PRPSPVHVAFERLEKELIGAGFGQYRATWQGEVAVKAAQDEPDQAAAAAESVR 169
QY 77 -ELQLSRVNHNPNIKLYGACLNLP--VCLVMEYARGGSLYNVLHGAEPLPYVTA----- 127
Db 170 REARLFAMLRHNPNIETLRGVCLQPHCLVLFARFGGALNRAAANAAPDPRAPCPRA 229
QY 128 ----AHAM-SWCLQCSQGVAYLHMQPKALIHRLDKPNLLLVAG-----GTVLKICD 175
Db 230 RRIPPHVLNVAVOIARGMLYLHHEAFVPIHLRDLKSSNILLLEKIEHDDICNKTKITD 289
QY 176 FGTACD-IQHTMTNKNKGAAMPAVEFGSNYSKCDVFSWGIILWEVITRRKPFDEIGG 234
Db 290 FGLAREWHRTTKMSTAGTYAMWAPAEVIKSSLSFGSGDILWSGVLLWELLTGEVYRGIDG 349
QY 235 PAFRIMWAVHNGTRPPLIKNLKPKIESLMTRCWSDPSQRPDS---MEEIVKIMTHLMRY 290
Db 350 LAVAYGVAWNLTL-PIPTCTPEPAKLKMECQWQDPIRPSFALLILEQLTAIEGAVNTE 408
QY 291 FPCA-----DELOYPCQVSDGQSNATSSTGFMFDIASTNSKSDTNM----EQVP 339
Db 409 MPQESFHSMDQDWKLEIQ-QMFDLRTKEKELRSREBELTRAALQKQSOEBLLKRREQOL 467
QY 340 ATN--DTIKRLESLLKNAQKQSSGRLSLGASHGSSVESLPTSEGRKMSADMSEIEA 397
Db 468 AEREIDVLER-ELNLIIFQLNQEKPKVKRKKGKFKRSLLK----LKQHRISLP-SDFOH 521
QY 398 RIAATGTGQPRRRSIQDLTVTGTGPGOVSSRSSSPSVRMITTSQPTSEK----- 447
Db 522 KITVQASPNLDKRSL-----NSSSSSPSPSPTMPLRAIQLTSDSNKTKWGNTVFRQ 576
```

```
QY 448 -----PTRGHPMTD-----DSTDTNG-----SDKSIPMAYLTLDHO-LOPLA-- 484
Db 577 EEPEDVKRNFKKGCTWGPNSIQMKDRITDCKERTRPLSDGNSPWSTLIINKQKTMPLASL 636
QY 485 -----PCPNKESMAVPEQHCCKMAQYMKVQTEIALLL-----QRKQELVAELDQDEKQ 534
Db 637 FVDQPGSCSEPKLSPDGL-HRKPKQIKLPSQAVIDLPLGKDAQRENPAEAE-SWEEAAS 694
QY 535 QNTSRLVQEHKKLLDENKSLSTYYQCKKQLLEVIRSOQQR 575
Db 695 ANAATVSTIE---MTPTNSLS-----RSPQRKK 718
```

RESULT 5

```
US-09-862-027-19
; Sequence 19, Application US/09862027
; Patent No. US20020142428A1
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin R.
; TITLE OF INVENTION: No. US20020142428A1el Kinases and Uses Thereof
; FILE REFERENCE: 35800/234862
; CURRENT APPLICATION NUMBER: US/09/862,027
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 09/345,473
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-862-027-19
```

```
Query Match 13.9%; Score 418.5; DB 10; Length 394;
Best Local Similarity 36.6%; Pred. No. 1.2e-18;
Matches 98; Conservative 47; Mismatches 100; Indels 23; Gaps 8;
```

```
QY 35 EIEVEVVGCGAFVVCCKAKRAKVAIK--QJESERKAFIVELRQ-----LSRVNHPN 88
Db 2 ELTLEIIGGGGKGVYRAFVIGDEVAVKARHDPEDISQTIENVRQBAKLFAMLKHPN 61
QY 89 IVKLYGACLNLP--VCLVMEYARGGSLYNVLHGAEPLPYVTAHAHMSWCLQCSQGVAYLHS 146
Db 62 IIALRGVCLKEPNLCLVMEFARGGPNRVLSGKRIPPDI----LVNWAYIARGMNYLHD 117
QY 147 MQPKALIHRLDKPNLLV-----AGTVLKICDFTACD-IQHTMTNKNKGAAMAP 198
Db 118 EAIVPIIHRDLKSSNILLILQKVENGDLSNKLKITDFGLAREWHRTTKMSAAGTYAMAP 177
QY 199 EVPEGSNYSKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMWAV-HNGTRPPLIKNLPK 257
Db 178 EVIRASWFSKGSVDWYGVLLWELLTGEVYFRGIDG--LRVAYGVAMNKLALPIPTCTPE 235
QY 258 PIESLMTRCWSDPSQRPDSMEIEIVKIMT 285
Db 236 PFAKLMEDCWNPDPSRPSFTNILDQLT 263
```

RESULT 6

```
US-09-947-199-2
; Sequence 2, Application US/09947199
; Patent No. US20020127684A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jayaseelan
; TITLE OF INVENTION: NOVEL CARC PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: THEREFOR
; CURRENT APPLICATION NUMBER: US/09/947,199
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: 60/111,938
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/291,839
; PRIOR FILING DATE: 1999-04-14
```

PRIOR APPLICATION NUMBER: 09/458,457
 PRIOR FILING DATE: 1999-12-10
 NUMBER OF SEQ ID NOS: 9
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 2
 LENGTH: 835
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-947-199-2

Query Match 13.9%; Score 418; DB 10; Length 835;
 Best Local Similarity 31.6%; Pred. No. 3e-18;
 Matches 115; Conservative 67; Mismatches 144; Indels 38; Gaps 13;

QY 31 IDYKEIEVEEVGRGAFGVYCKAKWRADVAIKOIE-----SESEKAFIVELRQLSRVN 85
 DB 458 IQLSEIEFHEIIGSGSFGKYVYKGRCKNKIYAIKRYRANTYCKSDVDFECREVSILCOLN 517
 QY 86 HPIVLYGACLN---PVCLVMEYAEAGSLVYVHLGAEPLPYTTAAHAMSWCLQCSQVA 142
 DB 518 HCVIQFVACLNDSQFAIVTYQYISGSLFSLH--EQKRILDLQSKLIIVADVAKGME 575
 QY 143 YLHSM-QPPALIHRLDKPPNLLVAGTVLKCDPFTACDIQT---HMTNKGSAAMA 197
 DB 576 YLHNLTOP--IHRDLNSHNLILYEDGHAV-VADFGESRFLQSLDEDNNTYQPGNLRWMA 632
 QY 198 PEYF-EGSNYSEKCDVFSMGIIIMWEYITRRKPFDEIGCFAPRIMAVHNGTRPPLIKNL 256
 DB 633 PEYFQCTYTTIADVFSTALCIWEITGEIPFAHKPAAAADAAVYH-IRPPIGYSI 691
 QY 257 KPISLMTKCSKDPSPQSRPMEIYKIMTHLMRYFPGADEPQYPCQYSDEGQSNATST 316
 DB 692 KPISLLIRGMWACPEGRPEFSEVVKLECL-----CNIELMSPA-----SSNS 737
 QY 317 GSGMDIASN-TSNKSDTMEQVPAINDTIKLESKLLKNQAKQSESGRLSLGASHSS 375
 DB 738 GSLSPSSSDCLVNRGPGRSHVALLRSRPELEYALNARSYALASQASQYS---SQGLS 794
 QY 376 VESL 379
 DB 795 LEBM 798

RESULT 7
 US-09-862-027-18
 Sequence 18, Application US/09862027
 Patent No. US20020142428A1
 GENERAL INFORMATION:
 APPLICANT: Hodge, Martin R.
 TITLE OF INVENTION: NO. US20020142428A1 Kinases and Uses Thereof
 FILE REFERENCE: 35800/234862
 CURRENT APPLICATION NUMBER: US/09/862,027
 CURRENT FILING DATE: 2001-05-21
 PRIOR APPLICATION NUMBER: US 09/345,473
 PRIOR FILING DATE: 1999-06-30
 NUMBER OF SEQ ID NOS: 82
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 18
 LENGTH: 328
 TYPE: PRT
 ORGANISM: C. elegans
 US-09-862-027-18

Query Match 13.8%; Score 416; DB 10; Length 328;
 Best Local Similarity 33.4%; Pred. No. 1.4e-18;
 Matches 99; Conservative 57; Mismatches 112; Indels 28; Gaps 11;
 QY 1 MSTASAASSSSSAGMIEAPSQVLFEEIDYKEIEVEEVGRGAFGVYCKAKWRADY 60
 DB 22 MSTSTSNESSTSSSNN-----DQRLVLPDIQRDDIQVGDHIGVTFGAVFSGNWTLPG 76
 QY 61 AIKQIESSEKATIVLRQ--LSRVNHPNIVKLYGACL---NPVCLVMEYAEAGSLVY 115

DB 77 SQRTI---ALKVYVLEKEAEILSKIRKNIIOFYICKATNDPFITYEAKSLDYDF 133
 QY 116 LHGAELPYTTAAHAMS-----WCLQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGT 169
 DB 134 IHSEESQSFASSSGNSFVUVKMASQIASGIOYLHYDAVDITIHRLDKSKNVVL-DKXL 192
 QY 170 VKICDPGTRACDIQTHMTNKK---GSAAMAPE-VPEGSNYSKCDVFSWGIIMWEVIT 225
 DB 193 VCKICDFGTSKDL-THSCAPSWGTAAMWSPMTIQSEGLTTATDVWSYGVVLEILSK 251
 QY 226 RKPDEIGCFAPRIMAV-HNGTRPPLIKNLKPISLMTKCSKDPSPQSRPMEI 280
 DB 252 EYVYKQYS--EPIIFMITQSGITTLALPSCAPLKLQMLNSCMKWTPKCRAMWQI 305

RESULT 8
 US-09-947-199-8
 Sequence 8, Application US/09947199
 Patent No. US20020127684A1
 GENERAL INFORMATION:
 APPLICANT: Raju, Jeyaseelan
 TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES
 FILE REFERENCE: MNI-06BCP2
 CURRENT APPLICATION NUMBER: US/09/947,199
 CURRENT FILING DATE: 2001-09-05
 PRIOR APPLICATION NUMBER: 60/111,938
 PRIOR FILING DATE: 1998-12-11
 PRIOR APPLICATION NUMBER: 09/291,839
 PRIOR FILING DATE: 1999-04-14
 PRIOR APPLICATION NUMBER: 09/458,457
 PRIOR FILING DATE: 1999-12-10
 NUMBER OF SEQ ID NOS: 9
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 8
 LENGTH: 835
 TYPE: PRT
 ORGANISM: Rattus norvegicus
 US-09-947-199-8

Query Match 13.7%; Score 414; DB 10; Length 835;
 Best Local Similarity 31.7%; Pred. No. 5.3e-18;
 Matches 120; Conservative 67; Mismatches 143; Indels 48; Gaps 15;

QY 20 EAPSQVLFEEIDYKEIEVEEVGRGAFGVYCKAKWRADVAIKOIE-----SESEKAF 74
 DB 451 ELPSRF---HLQSEIEFHEIIGSGSFGKYVYKGRCKNKIYAIKRYRANTYCKSDVDF 506
 QY 75 IVELRQLSRVNHPIVLYGACL---NPVCLVMEYAEAGSLVYVHLGAEPLPYTTAAHAM 131
 DB 507 CREVSILCOLNHPCVVQPVGACLDDEPQFAIVTYQYISGSLFSLH--EQKRILDLQSKL 564
 QY 132 SWCLQCSQGVAYLHSM-QPKALIHRLDKPPNLLVAGTVLKCDPFTACDIQT---HM 186
 DB 565 IIVADVAKGMEYHSLQGP--IHRDLNSHNLILYEDGHAV-VADFGESRFLQSLDEDNM 621
 QY 187 TNNKGAAMAPPEYF-EGSNYSEKCDVFSMGIIIMWEYITRRKPFDEIGCFAPRIMAVHN 245
 DB 622 TKQPGNLRWMAPEVFTQCTRTYIKADVFSYSLCIWEITGEIPFAHKPAAAADAAVYH 681
 QY 246 GTRPPLIKNLKPISLMTKCSKDPSPQSRPMEIYKIMTHLMRYFPGADEPQYPCQYS 305
 DB 682 -IRPPIGYSIKPISSILIRGMNACPEGRPEFSEVVKLECL-----CNVELMSPA--- 732
 QY 306 DEGQSNATSTGSPMDIASN-TSNKSDTMEQVPAINDTIKLESKU-LKNQAKQSESG 364
 DB 733 -----GSNSGSL-----SPSSSDCLLSRGPGRSHVALLRSRPELEYALNARSYAG 780
 QY 365 RSLGLASH---GSSVESL 379
 DB 781 WSQSVGTHSNPGLSLEEM 798

```

; GENERAL INFORMATION:
; APPLICANT: Dedhar, Shoukat
; APPLICANT: Hannigan, Greg
; TITLE OF INVENTION: Intergrin-Linked Kinase and its Uses
; FILE REFERENCE: KIN-2CON
; CURRENT APPLICATION NUMBER: US/09/840,704
; CURRENT FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: 09/566,906
; PRIOR FILING DATE: 2000-05-09
; PRIOR APPLICATION NUMBER: US08/752,345
; PRIOR FILING DATE: 1996-11-19
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 263
; TYPE: PRT
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: Other
; LOCATION: (1)...(263)
US-09-840-704-5

Query Match          13.0%; Score 393; DB 10; Length 263;
Best Local Similarity 33.8%; Pred. No. 2.6e-17;
Matches 90; Conservative 57; Mismatches 107; Indels 12; Gaps

Qy 31 IDYKEIEVEEVGAGFYGVCKAKRWADVAIK--QIESESER-KAFIVELRQLSRVNH 86
Db 1 IPWCDLNIKEIKIGAGSGFTVHRAEWHGSDVAVKILMEQDPHAE RVNVEFLREVAIMKELRH 60
Qy 87 PNIVKLYGACLNLP--VCLVMYEAGGSLYLVNLHGAEPLPYTTAAHAMSWCLOCSQGVAYL 144
Db 61 PNIVLFGAVTQPPNLSIVTEYLSRGSLYRLHLKSGAREQLDERRRLSMAYDVAKGMNYL 120
Qy 145 HSMOPKALIHRLDKPNNLLVAGGTVLVKICDFG-TACDIOHTHTNNK--GSAAWMAPEVF 261
Db 121 HNRNP-PIVHRDLKSPNLLVDKKYTV-KVCDGLSLRKAKSTFLSSKSAAGTPEWMAPEVL 178
Qy 202 EGSNYSKCDVFSWGIILMEVITERRKPFDEIGGPAFRIMWAVHNGTRPPLIKNLPKPIES 261
Db 179 RDEPSNEKSDVYSFGVILWELATLQDPWGNL-NPAQVVAAGVGFCKRLEIPRNLNPFQVAA 237
Qy 262 LMTFCWSKDPQRSMEIEIVKIMTHL 287
Db 238 IIEGWTNEPWRKPSFATIMDLLRPL 263

RESULT 11
US-09-904-389-2
; Sequence 2, Application US/09904389
; Patent No. US20020129404A1
; GENERAL INFORMATION:
; APPLICANT: Clendennen, Stephanie K.
; APPLICANT: Schuster, Debra K.
; TITLE OF INVENTION: CTRI HOMOLOGUE FROM MELON
; FILE REFERENCE: 4257-0029.30
; CURRENT APPLICATION NUMBER: US/09/904,389
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: US 60/218,307
; PRIOR FILING DATE: 2000-07-14
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 850
; TYPE: PRT
; ORGANISM: Cucumis melo
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (154)...(154)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-904-389-2

Query Match          13.0%; Score 392; DB 10; Length 850;

```


Job time : 119.014 secs

Matches 105; Conservative 42; Mismatches 81; Indels 67; Gaps 13;

[illegible]

RESULT 15

US-09-842-582-4

; Sequence 4, Application US/09842582

; Patent No. US20020155570A1

; GENERAL INFORMATION:

; APPLICANT: Millennium Pharmaceuticals, Inc.

; APPLICANT: Meyers, Rachel

; TITLE OF INVENTION: 2246, NOVEL PROTEIN KINASE MOLECULES AND

; TITLE OF INVENTION: USES THEREFOR

; FILE REFERENCE: 38155-20054.00

; CURRENT APPLICATION NUMBER: US/09/842,582

; CURRENT FILING DATE: 2001-04-25

; PRIOR APPLICATION NUMBER: US 60/199,391
 ; PRIOR APPLICATION NUMBER: US 60/199,391

;; PRIOR FILING DATE: 2000-04-25

; NUMBER OF SEQ ID NOS: 11
SOFTWARE: Easel v2.2.0 from WI

```

; SOFTWARE: FASTSEQ FOR WINDOWS VERSION 4.0
: SEQ ID NO 4

```

```

; SEQ ID NO 4
: LENGTH: 279

```

```

;
: TYPE: PPT
LENGTH: 278

```

; TYPE: PRI
; ORGANTCM:

```

; ORGANISM: Artificial sequence
: FEATURE:

```

FEATURE: OTHER TAX

OTHER INFORMATION: Consensus amino acid
UIS-09-842-582-4

US-09-842-582-4

Query Match 11.9%: Score 359: DB 9: Length 278:

Best Local Similarity 35.6%; Pred. No. 3.2e-15;

Sequence Alignment	Matches	Conservative	Mismatches	Indels	Gaps
100% Identity	105	42	82	66	13

Qy	37	EVEEVVGEAGVVCAKWR--AKOVAIKQTESSEERKAFIVELPOLSRVNHPIVKLYGA	95
Db	2	ELLEKLGEGSGKYVKAHKTKGIKVAKILKKES---LSUREITQILKRLSHPIVRLLVG	58
Qy	96	CLNP----VCLWVEAEGGSLNVLHGAEPLPYPTAAHAMSCWCSQGVAYLHSMQPKAL	152
Db	59	FEDTDDHLYLWMEYEGGDLFDYLRRNGPL---SEKEAKIALQILRGLEYLHS---	112
Qy	153	IHRDLKPNPILLVAGGTIVLKICDGTGACDIOTHTNNKGSAAW--MAPEV-FRGSNYSEK	209
Db	113	VHRDLKPNPILLDNGTV-KIADFGLA-RLLEKILTTFTGT-PWYMAAPEVILEGRGYSSK	169
Qy	210	CDVFSWGIIHWEVI-----TRRKPFE	231
Db	170	VDVWSLGVILYELLTGGPLPGADLPAGTGDGVDQLIIVLVLKPSDELPKTRIDPLEE	229
Qy	232	IGGPAFRIMVAHNGTRPPLIKNLPKPIESLMTFCWSKQSPQSRPSMEIIVKIMTH	286
Db	230	L-----FR-----KKRRLLPSPNGSEELKDLLKKLKNKQSPKRGSAATAKEILNH	275

Search completed: December 9, 2002, 23:04:25

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 9, 2002, 22:58:34 ; Search time 3633.32 Seconds
(without alignments)
4637.778 Million cell updates/sec

Title: US-09-830-144-2

Perfect score: 3014

Sequence: 1 MSTAASASSSSSSAGEMIE.....QCKQLEVRSSQQKRGQTS 579

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-Q=/cgn2.1/USPTO.spool/US09830144/runat_04122002.141353.2252/app_query.fasta_1.1422
-DB=GenEmbl -OPMT=fastap -SURFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DUALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09830144 @cgn_1.1 5173 @runat_04122002.141353.2252 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vl.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.man.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	3006	99.7	2757	9	BC017715	BC017715 Homo sapi
2	3006	99.7	2769	6	AX377912	AX377912 Sequence
3	3006	99.7	2769	9	AB009356	AB009356 Homo sapi
4	3006	99.7	2785	6	E38397	E38397 NF-kappa B
5	2986	99.1	2443	10	MUSTAK1	D76446 Mouse mRNA
6	2982.5	99.0	2850	9	AB009357	AB009357 Homo sapi
7	2982.5	99.0	2866	6	E38398	E38398 NF-kappa B
8	2963.5	98.3	3107	10	BC006665	BC006665 Mus muscu
9	2756.5	91.5	1745	9	AF218074	AF218074 Homo sapi
10	2746	91.1	1704	6	E38399	E38399 NF-kappa B
11	2729	90.5	1705	9	AB009358	AB009358 Homo sapi
12	2692	89.3	135147	2	AC114407	AC114407 Mus muscu
13	2638.5	87.5	2812	5	XLU92030	U92030 Xenopus lae
14	1476	49.0	3482	9	HSM800550	AL050393 Homo sapi
15	971	32.2	3349	3	AF159466	AF159466 Drosophill
16	971	32.2	3366	3	AY051933	AY051933 Drosophill
17	910	30.2	2213	9	AK055901	AK055901 Homo sapi
18	702.5	23.3	32564	2	AC014558	AC014558 Drosophill
19	702.5	23.3	164942	3	AC011758	AC011758 Drosophill
c	702.5	23.3	302303	3	AE003571	AE003571 Drosophill
c	502	16.7	3138	9	HSM8TMR	Z48615 H.sapiens M
21	493.5	16.4	3454	6	AX337846	AX337846 Sequence
22	493.5	16.4	3454	9	HARNAMUK2	X90846 H.sapiens m
23	490	16.3	2120	6	AR119790	AR119790 Sequence
24	490	16.3	2120	6	AR126750	AR126750 Sequence
25	490	16.3	2120	6	AR128910	AR128910 Sequence
26	490	16.3	2120	6	AR130841	AR130841 Sequence
27	490	16.3	2120	6	AR138886	AR138886 Sequence
28	490	16.3	2120	6	AR141354	AR141354 Sequence
29	482	16.0	1370	9	AF325454	AF325454 Homo sapi
30	482	16.0	1571	9	BC001401	BC001401 Homo sapi
31	482	16.0	2251	9	AK056310	AK056310 Homo sapi
32	482	16.0	7195	9	AF480462	AF480462 Homo sapi
33	481	16.0	1365	6	AR119791	AR119791 Sequence
34	481	16.0	1365	6	AR126751	AR126751 Sequence
35	481	16.0	1365	6	AR128911	AR128911 Sequence
36	481	16.0	1365	6	AR130842	AR130842 Sequence
37	481	16.0	1365	6	AR138887	AR138887 Sequence
38	481	16.0	1365	6	AR141355	AR141355 Sequence
39	481	16.0	1368	9	AB049734	AB049734 Homo sapi
40	481	16.0	1429	10	AB049732	AB049732 Mus muscu
41	481	16.0	3931	9	AF251442	AF251442 Homo sapi
42	480	15.9	8125	3	AF481923	AF481923 Dictyoste
43	474	15.3	3531	9	HSU07747	U07747 Human SH3 d
44	462	15.3	3558	6	AX399680	AX399680 Sequence
45	462	15.3	3558	6	AX399680	AX399680 Sequence

ALIGNMENTS

RESULT 1

BC017715	2757 bp	mrna	linear	PRI 06-DEC-2001
LOCUS				
DEFINITION	BC017715	mitogen-activated protein kinase kinase 7,		
ACCESSION	BC017715	Clone MGC:21263 IMAGE:3906837, mRNA, complete cds.		
VERSION	BC017715.1	GI:17389342		
KEYWORDS	MGC.			
SOURCE	Homo sapiens.			
ORGANISM	Homo sapiens.			
REFERENCE	Eukaryote: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
AUTHORS	1 (bases 1 to 2757)			
TITLE	Strausberg,R.			
JOURNAL	Direct Submission			
REMARK	Submitted (03-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA			
COMMENT	NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk Email: cgapds-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305 Web site: http://www.sngc.stanford.edu Contact: (Dickson, Mark) mcdpaxil@stanford.edu Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.			
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Series: IRAX Plate: 22 Row: 1 Column: 18 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4507360.			
SOURCE	Location/Qualifiers			
	1..2757			
	/organism="Homo sapiens"			
	/db_xref="locusID:6885"			
	/db_xref="taxon:9606"			
	/clone="MGC:21263 IMAGE:3906837"			
	/tissue.type="uterus", leiomyosarcoma"			
	/clone_id="NIH_MGC_71"			
	/lab_host="DH10B"			
	/note="vector: pCMV-SPORT6"			
	198..1937			
	/codon_start=1			
	/product="mitogen-activated protein kinase kinase 7"			
	/protein_id="AAH17715.1"			
	/db_xref="GI:17389343"			
	/translation="MSTASAASSSSSSSAGEMTEAPQVYLFPEIDYKEIFVEVGR			
	GAFVCAKAKRKAQVAIKQIESERKAFIVELRQUSRVNHNIVIKYGCALNPVCL			
	VMEYAGSLVNLVHGADPLPYITLAHMSQCLQSGVAFLHSMQKALIHRLKPP			
	NLLVAFGSLVLIKIDFGFACDIOTHTNNKSSAAMAPVEGNSYSPKCDVFSWGII			
	LMEVTRRKPRPEIGGPAIRFVAHNGTRPPLIKNLPIESLMTKMSKDPQSPRS			
	MEETVKNITHLMRYFPQADPELVQPCVSDGQASNTSGSPMDIASTNTSNKSDIN			
	MEQVATPNTDIKRLSKLLKQAKQSSGRSLGASGVSESLPIPSBEKMSADM			
	SEIERIAATNGQPPRRRSIQDITVTSTETPEQVSSRSFSVRKITITSGTSEKPTR			
	SHHPMTDDSTNGSDNSIPMAVLIILDLQOLPACPKSKSMVFEQHCMAQEVNKM			
	VTETALLORKEQLVAELIDDEKQOQNTSLVGRHKLLDENKLSLTVYQOCCKQLE			
	VTRSQOQKQKQGS"			
BASE COUNT	819 a 566 c 647 g 725 t			
ORIGIN				
Alignment Scores:				
Pred. No.:	4.88e-177	Length:	2757	
Score:	3006.00	Matches:	578	
Percent Similarity:	99.83%	Conservative:	0	
Best Local Similarity:	99.83%	Mismatches:	1	
Query Match:	99.73%	Indels:	0	
DB:	9	Gaps:	0	

[illegible]

Db 343 GCTATTAAACAATAGAAAGTGAATCTGAGAGAAAGCGTTATTGTAGAGCTTCGGCAG 402
 QY LeuSerArgValAsnHisProAsnIleValLeuTyrGlyAlaCysLeuAsnProVal 100
 Db 403 TTATCCCTGCTGAACCACTTAATATGTAAAGCTTTATGAGCTGTGATTCAGTG 462
 QY CysLeuValMetGluTyrAlaGlyGlySerLeuTyrAsnValLeuHisGlyAlaGlu 120
 Db 463 TGTCTTGATGAGATATGCTGAAAGGGGCTTTATATATATGCTGCATGGCTGAA 522
 QY ProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeuGlnCysSerGlnGly 140
 Db 523 CCATTGCCATATTATAGCTGTCGCCCAACCAATGAGTTGTGTTTAAAGTGTTCACAAAGA 582
 QY ValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysProPro 160
 Db 583 GTGGCTTATCTTCACAGATGCAACCCAAAGCGCTTAATTCACAGGAGCTTAACACACA 642
 QY AsnLeuLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCys 180
 Db 643 AACTTACTGCTGCTGTGACAGGGGGGACAGTCTTAATAATTTGTGATTTGTGATACAGCTGT 702
 QY AspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaIleTyrMetAlaProGluVal 200
 Db 703 GACATTCAGACACACATGACCAATTAACAAGGAGTGTGCTGTGATGCGACCTGAAAGTT 762
 QY PheGluGlySerAsnTyrSerGlyLysCysAspValPheSerTyrGlyIleIleLeuTyr 220
 Db 763 TTTGAAGGAGTAAATTACAGTAAAAATGTGACGTTCACTGGGGATTAATTTCTTGG 822
 QY GluValIleLeuThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMet 240
 Db 823 GAAGTGATTAACCGCTGGAAACCTTTGATGATGATGTTGGGCCCAAGCTTCGATCATCG 882
 QY TrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGlu 260
 Db 883 TGGGCTGTTCAATATGATGCTGACACCATGATTAATAATTTACATTAAGCCCATTTAG 942
 QY SerLeuMetThrArgCysTyrPheLysAspProSerGlnArgProSerMetGluGluIle 280
 Db 943 AGCGTATGACTCGTGTGTGTCTTAAGATCCTTCCACAGCGCTTCAATGAGGAAAT 1002
 QY ValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyr 300
 Db 1003 GTGAAATATATGACTGACTGATGCGGATCTTCCACAGACAGATGACCATTAAGAT 1062
 QY ProCysGlnTyrSerAspGluGlyLysSerAsnSerAlaThrSerThrLysSerPheMet 320
 Db 1063 CCTTGCAGATTCATGATGAGACAGACAGACACTTCCACACAGTACAGGCTCATTCATG 1122
 QY AspIleAlaSerThrAsnThrSerAsnLysSerAspThrAsnMetGluGlnValProAla 340
 Db 1123 GACATTGCTTCTCAATAATGAGATAACAAAGTACACATTAATGAGCAAGTCTCTGCC 1182
 QY ThrAsnAspThrIleLysArgLeuGlyLysSerLeuLeuLysAsnAlaLysGlnGln 360
 Db 1183 ACAAAATGATATCTATTAAAGCGCTTAAGATCAAAATGTTGTAATAATGAGCAAAACAG 1242
 QY SerGluSerGlyArgLeuSerLeuGlyAlaSerHisGlySerSerValGluSerLeuPro 380
 Db 1243 AGTGAAATCTGAGCTTTAAAGCTTGGAGAGCTCCCGTGGAGAGCTGTGAGAGCTTGCC 1302
 QY ProThrSerGluGlyLysArgMetSerAlaAspMetSerGluIleGluAlaArgIleAla 400
 Db 1303 CCAACCTCTGAGGCAAGAGAGATGAGTGCATGATCTGAAATAGAAAGCTAGATCGCC 1362
 QY AlaThrThrGlyAsnGlyGlnProArgArgArgSerIleGlnAspLeuThrValThrGly 420
 Db 1363 GCAACCCACAGGCAACGAGACGCAAGCGTAGATCCATCCAAAGCTTGACTGTAACCTGGA 1422
 QY ThrGluProGlyGlnValSerSerArgSerSerProSerValArgMetIleThrThr 440
 Db 1423 ACAGAACTGGTCAAGTGAGCAGTAGTCAATCCAGTCCAGTGTCAAGAAATGATTACTACC 1482

QY 441 SerGlyProThrSerGluLysProThrArgSerHisProTyrThrProAspAspSerThr 460
 Db 1483 TCAGGACCAACCTCAGAAAAGCAACTCGAAGTCAATCCATGAGACCCCTGATGATTCACA 1542
 QY 461 AspThrAsnGlySerAspAsnSerIleProMetAlaTyrLeuThrLeuAspHisGlnLeu 480
 Db 1543 GATACCAATGAGATCAGATTAATCCATCCCAATGCTTATCTTCACTGATCACCACATA 1602
 QY 481 GlnProLeuAlaProCysProAsnSerLysGlySerMetAlaValPheGluGlnHisCys 500
 Db 1603 CAGCTCTGACACCGTGGCCCAACCAAGATCATAGGACGAGTGTGACAGCATTTGT 1662
 QY 501 LysMetAlaGlnGluTyrMetLysValGlnTyrGluIleAlaLeuLeuGlnArgLys 520
 Db 1663 AAAATGCAACAAGATATATGAAAGTTCAACAGAAATTCGATTTGATTAAGAGAAAG 1722
 QY 521 GlnGluLeuValAlaGluLeuAspGlnAspGlyLysAspGlnGlnAsnThrSerArgLeu 540
 Db 1723 CAAGAACTAGTTGACAGAACTGACACGAGATGAAGAGACACCAAAATACATCTGCCTG 1782
 QY 541 ValGlnGlnHisLysLysLeuLeuAspGluAsnLysSerLeuSerThrTyrTyrGlnGln 560
 Db 1783 GTACAGAACATTAAGCTTTAGTGAATAACAAAGCTTTCACTTACTTACACAGCAA 1842
 QY 561 CysLysLysGlnLeuGluValIleArgSerGlnGlnGlnLysArgGlnGlyThrSer 579
 Db 1843 TGCAAAAACAATGACATGATCAGAGTCAAGACAGCAAGAAACGACGACTTCA 1899

RESULT 5
 MUSTAKI
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

1
 MUSTAKI
 Mouse mRNA for TAK1 (TGF-beta-activated kinase), complete cds.
 D76446.1 GI:1167505
 TAK1, TGF-beta-activated kinase; protein kinase.
 Mus musculus
 Mus musculus
 Mus musculus
 Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (sites)
 Yamaguchi, K., Shitake, K., Shibuya, H., Irie, K., Oishi, I., Ueno, N.,
 Taniguchi, T., Nishida, E. and Matsumoto, K.
 Identification of a member of the MAPKs family as a potential
 mediator of TGF-beta signal transduction
 Science 270 (5244), 2008-2011 (1995)

JOURNAL
 MEDLINE
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

96123277
 2 (bases 1 to 2443)
 Matsumoto, K.
 Direct Submission
 Submitted (18-OCT-1995) Kunihiko Matsumoto, Faculty of Science,
 Nagoya University, Department of Molecular Biology, Furou-chou,
 Chikusa-ku, Nagoya, Aichi 464-01, Japan
 (E-mail: 944177@nuc.cc.nagoya-u.ac.jp, Tel: 052-789-3000,
 Fax: 052-789-3001)

FEATURES
 source
 location/Qualifiers
 1..2443
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 157..1896
 /codon_start=1
 /product="TAK1 (TGF-beta-activated kinase)"
 /protein_id="BA01184.1"
 /db_xref="GI:1167506"
 /translation="MSTASAASSSSSSASEMTEAPSOVLENEIDYKEIYEVEVGR
 GARGVCKAKRKADVAIKQIBESERKAFYELRQLSKVNHNYIKYLGACINPCL
 VMEYABGSLYNVLHGAEPDLYTAAHAWSCQSGVAVLHWPKALIHRLKPP
 NLLVAGTVLKI CDFTGACDIQTHMTNNKSGAAWMAPEVFGSNSEKDFVSWGI
 LMEVTRKRPDEIGGPAFRIIMAVHNGRPPILKILPKPIESLMTKCSKDPSPS
 MERVKTMTLMVYFPGADBPLOVPOYSEGGNSATSGMDIADTNTSNKSDTN
 MEQVPATNDITKLESKLKNQKQOSEGRSLISLGRSISVSLPPTSEGRKSMAD
 SEIETARIVATANGQPRRSIOLVITVTEBPQVSSRSSPSYRMTTTSGRSEKAR
 SHPWPBDSYDTNGSDNSIPMAVLTLDHQLAPCPNSKESMAVEHCKMAQETWK

VOTEIALLLQKQELVAELDDQDKQNTSRLVQBHKLLDENKSLSTYYQCKKQLE
VIRSOQKQKQGS"

BASE COUNT 669 a 568 c 647 g 559 t
ORIGIN

Alignment Scores:

Pred. No.: 7,3e-176 Length: 2443
Score: 2986.00 Matches: 574
Percent Similarity: 99.14% Conservative: 0
Best Local Similarity: 99.14% Mismatches: 5
Query Match: 99.07% Indels: 0
DB: 10 Gaps: 0

US-09-830-144-2 (1-579) x MUSTAK1 (1-2443)

Qy 1 MetSerThrAlaSerAlaSerSerSerSerSerSerSerAlaGlyGluMetIleGlu 20
Db 157 ATGTCGACAGCCTCGCGCGCTGCTCCTCCTCGCTTCTGCCAGTGAGATCGAA 216
Qy 21 AlaProSerGlnValLeuAsnPheGluIleAspTyrIysGluIleGluValGlu 40
Db 217 GCGCGCTGCGAGGTCCTGAACCTCGAAGAGATCGACTACAAAGGAGATCGAGGTGGAAG 276
Qy 41 ValValGlyArgGlyAlaPheGlyValValCysIysAlaIysTyrArgAlaIysAspVal 60
Db 277 GTTGTGGAAGAGAGCTTTTGGAGTAGTTTGCAGAGCTTAAGTGGAGAGCAAAAGATGTC 336
Qy 61 AlaIleIysGlnIleGluSerGluSerGluArgIysAlaPheIleValGluIleArgGln 80
Db 337 GCTATTAAACAGATAGAAAGTAGTCTGAGAGGAAGCTTTTCAATTGGAGCTCGGCGAG 396
Qy 81 LeuSerArgValAsnHisProAsnIleValIysLeuTyrGlyAlaCysLeuAsnProVal 100
Db 397 TTGTCGCTGTGAACCATCTTAACATTTGTCAGTTGTACGAGACCTTGCCTGAATCCAGTA 456
Qy 101 CysLeuValMetGluTyrAlaGluGlyGlySerIleTyrAsnValLeuHisGlyAlaGlu 120
Db 457 TGTCTTGTGTAGTAATATGACAGAGGGGGCTCATTTGATATTAATGTGCTGCTGTGAA 516
Qy 121 ProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeuGlnCysSerGlnGly 140
Db 517 CCATTGCTTACTACATGCTGCTCATGCCATGAGCTGTGTGTTTACAGTGTCCCAAGGA 576
Qy 141 ValAlaTyrLeuHisSerMetGlnProIysAlaLeuIleHisArgAspLeuIysProPro 160
Db 577 GTGGCTTACCTGCACAGCATCAGCCCAAGCGCTGATTACAGGAGACCTCAAGACCTCCA 636
Qy 161 AsnLeuLeuValAlaGlyGlyThrValLeuIysIleCysAspPheGlyThrAlaCys 180
Db 637 AACTTGTCTGTGTGAGGAGGACAGTTCTAAAAATCTGCGATTTTGTGTACAGCTTGT 696
Qy 181 AspIleGlnThrHisMetThrAsnAsnIysGlySerAlaAlaTyrMetAlaProGluVal 200
Db 697 GACATCCAAACACACATGACCAATAATAAAGGAGTGCTGCTTGGATGGCGCTGAAGTG 756
Qy 201 PheGluGlySerAsnTyrSerGluIysCysAspValPheSerTyrGlyIleIleLeuTyr 220
Db 757 TTTGAAGTAGCAATATACAGTAAAGGTGTGATGCTTTCAGCTGGGGTATATCCCTCTGG 816
Qy 221 GluValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMet 240
Db 817 GAAGTGATAACACGCCGAACCCCTTCGATCAGATCGGTGGCCAGCTTTCAGAAATCATG 876
Qy 241 TrpAlaValHisAsnGlyThrArgProProLeuIleIysAsnLeuProIysProIleGlu 260
Db 877 TGGGCTGTTCATAATGGCACTCGACCACTGATCAAAAAATTTTACCCTAAGCCCATTTGAG 936
Qy 261 SerLeuMetThrArgCysTyrSerIysAspProSerGlnArgProSerMetGluGluIle 280
Db 937 ACCTTGATGACACAGCTGTGTGTCTAAGACCATCTCAGCGCCCTTCATGTGAGGAAT 996
Qy 281 ValIysIleMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyr 300

Db 997 GTGAAATAATGACTCACTTGATGCGGTACTTCCAGGAGCGGATGAGCCATTTACAGTAT 1056
Qy 301 ProCysGlnTyrSerAspGluGlyGlnSerAsnSerAlaThrSerThrGlySerPheMet 320
Db 1057 CCTTGTCACTACTCTGATGAAGGCGAGCAACTCAGCCACCAGCACAGGCTCGTTTCATG 1116
Qy 321 AspIleAlaSerThrAsnThrSerAsnIysSerAspThrAsnMetGluGlnValProIle 340
Db 1117 GACATTGCTTCTACAAATACCAAGTAAATAAAGTGACACAAATATGAAACAGGTTCCCTGCC 1176
Qy 341 ThrAsnAspThrIleIysArgLeuGluSerIysLeuLeuIysAsnGlnAlaIysGlnGln 360
Db 1177 ACAACGACACTATTAAACGCTTGGAGTCAAACTGTTGAAAAACCGAGCAAAAGCAACAG 1236
Qy 361 SerGluSerGlyArgLeuSerIleuGlyAlaSerHisGlySerSerValGluSerLeuPro 380
Db 1237 AGTGAATCTTGAGCGCTGAGCTTGGAGCCTCTCGTGGGAGCAGTGTGGAGAGCTTGGCC 1296
Qy 381 ProThrSerGluGlyIysArgMetSerAlaAspMetSerGluIleGluAlaArgIleAla 400
Db 1297 CCACCTCCGAGGGCGAAGGATGAGTGTGACATGTCTGAATATAGAAGCCAGGATCGTG 1356
Qy 401 AlaThrThrGlyAsnGlyGlnProArgArgSerIleGlnAspLeuThrValThrGly 420
Db 1357 GCGACTGCAAGTAAACGGCAACCAAGCGTAGATCCATCCAAGACTTGACTGTTACTGGG 1416
Qy 421 ThrGluProGlyGlnValSerSerArgSerSerProSerValArgMetIleThrThr 440
Db 1417 ACAGAACTCGTCAGGTGAGCAGCGGTATCCAGCCCTAGTGTGAGATGATCACTACC 1476
Qy 441 SerGlyProThrSerGluIysProThrArgSerHisProTyrThrProAspAspSerThr 460
Db 1477 TCAGGACCAACCTCAGAGAAGCCAGCTCGCAGTCAACCCATGACCCCTGATGATCCACA 1536
Qy 461 AspThrAsnGlySerAspAsnSerIleProMetAlaTyrLeuThrIleuAspHisGlnLeu 480
Db 1537 GACACCAATGGCTCAGATAACTCCATCCCAATGGCGTATCTTACACTGGATCACCAGCTA 1596
Qy 481 GlnProLeuAlaProCysProAsnSerIysGluSerMetAlaValPheGluGlnHisCys 500
Db 1597 CAGCTCTAGCGCGTGGCCCAACTCCAAAGAAATCCATGGCAGTGTTCGAACAGCACTGT 1656
Qy 501 LysMetAlaGlnGluTyrMetIysValGlnThrGluIleAlaIleuLeuGlnArgIys 520
Db 1657 AAAATGCCACAGAGTATATGAAAGTTCAACCGAAATCGCATTTGTACTACAGAAAG 1716
Qy 521 GlnGluLeuValAlaGluLeuAspGlnAspGluIysAspGlnGlnAsnThrSerArgLeu 540
Db 1717 CAAGAACTAGTTGCAGAAATTTGGACAGGATGAAAAGGACGACGCAAAATACATCTGCTG 1776
Qy 541 ValGlnGluHisIysIysLeuLeuAspGluAsnIysSerIysSerThrTyrTyrGlnGln 560
Db 1777 GTACAGGAACATAAAAAGCTTTTAGATGAAAAACAAAGAGCTTTTCTACTATTATACAGCAA 1836
Qy 561 CysIysIysGlnLeuGluValIleArgSerGlnGlnIysArgGlnGlyThrSer 579
Db 1837 TCGAAAAACAACCTAGAGTTCATCAGAACCAACAGCAAAACGACAGGCACTTCA 1893

RESULT 6

AB009357

LOCUS

DEFINITION

AB009357

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (sites)

REFERENCE

AUTHORS

TITLE

AB009357 Homo sapiens mRNA for TGF-beta activated kinase 1b, complete cds.

AB009357.1 GI:2924625

TAklb; TGF-beta activated kinase 1b.

Homo sapiens lung cDNA to mRNA, clone lib: Lambda gt11

clone: pBSTAK1b.

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (sites)

REFERENCE

AUTHORS

TITLE

TGF-beta-activated kinase 1 stimulates NF-kappa B activation by an

QY 474 LeuThrLeuAspHisGlnLeuGlnProLeuAlaProCysProAsnSerLysGluSerMet 493
 |||||
 Db 1663 CTTACTAGTACCACTACAGCCTCTAGCAGCGCTGCCAACTCCAAAGAAATCTATG 1722
 |||||
 QY 494 AlaValPheGluGlnHisCysLysMetAlaGlnGluTyrMetLysValGlnThrGluIle 513
 |||||
 Db 1723 GCAGTGTTCGAACAGCATTTGAAATGGCAAGAAATATATGAAAGTTTCAACAGAAAT 1782
 |||||
 QY 514 AlaLeuLeuGlnArgLysGlnGluLeuValAlaGluLeuAspGlnAspGluLysAsp 533
 |||||
 Db 1783 GCATTGTATTACAGAAAGCAGAACTAGTTGCGAACTGGACAGGATGAAAGGAC 1842
 |||||
 QY 534 GlnGlnAsnThrSerArgLeuValGlnGluHisLysLysLeuLeuAspGluAsnLysSer 553
 |||||
 Db 1843 CAGCAAAATACATCTCGCTGTTACAGGAACATAAAAGCTTTTACATGAAACAAAGC 1902
 |||||
 QY 554 LeuSerThrTyrGlnGlnCysLysLysGlnLeuGluValIleArgSerGlnGlnGln 573
 |||||
 Db 1903 CTTTCTACTTACTACAGCAATGCAAAACAACTAGAGTCAATCAGAGTCATCAGAGCAG 1962
 |||||
 QY 574 LysArgGlnGlyThrSer 579
 |||||
 Db 1963 AAACGACAGGCACCTTCA 1980
 |||||
 RESULT 7
 E38398
 LOCUS E38398 2866 bp DNA linear PAT 31-JAN-2002
 DEFINITION NF-kappa B activation inhibitory drug targeting TAK1 and method for
 identifying the same.
 E38398
 ACCESSION E38398.1 GI:18626978
 VERSION JP 2000197500-A/4.
 KEYWORDS unclassified.
 SOURCE unclassified.
 ORGANISM unclassified.
 REFERENCE 1 (bases 1 to 2866)
 AUTHORS Sugita,N., Sakurai,H., Kageyama,N. and Hasegawa,H.
 TITLE NF-kappa B activation inhibitory drug targeting TAK1 and method for
 identifying the same
 JOURNAL Patent: JP 2000197500-A 4 18-JUL-2000;
 TANAHE SEIYAKU CO LTD
 COMMENT OS Unidentified
 PN JP 2000197500-A/4
 PD 18-JUL-2000
 PE 04-FEB-1999 JP 1999026803
 PR
 PI NAOHISA SUGITA,HIROAKI SAKURAI,NORIKO KAGEYAMA,PI HIROSHI
 HASEGAWA
 PC C12Q1/48,A61K31/00,A61K31/00,A61K45/00,C12N5/10,C12N9/99,PC
 C12Q1/02
 PC G01N33/15,G01N33/50,G01N33/566/C12N15/09,C12Q1/68,C12N15/09,
 C12R1:911,
 PC C12N5/00,C12N15/00,C12N15/00,C12R1:91
 CC Strandedness: Double;
 CC Topology: Linear;
 FH Key Location/Qualifiers
 FT source 1..2866 /organism='Unidentified'.
 FT Location/Qualifiers
 source 1..2866
 /organism='unidentified'
 /db_xref='taxon:32644'
 BASE COUNT 848 a 588 c 656 g 774 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,45e-175 Length: 2866
 Score: 2982.50 Matches: 578
 Percent Similarity: 95.38% Conservative: 0
 Best Local Similarity: 95.38% Mismatches: 1
 Query Match: 98.95% Indels: 27
 DB: 6 Gaps: 1

US-09-830-144-2 (1-579) x E38398 (1-2866)
 QY 1 MetSerThrAlaSerAlaAlaSerSerSerSerSerSerAlaGlyGluMetIleGlu 20
 |||||
 Db 163 ATGCTACAGCCTTCGCGCT 222
 |||||
 QY 21 AlaProSerGlnValLeuAsnPheGluGluIleAspTyrLysGluIleGluValGluGlu 40
 |||||
 Db 223 GCCCTTCCCAGGCTCTCAACTTTGAAGAGATCGACTACAAGGAGATCGAGGTGGAAGAG 282
 |||||
 QY 41 ValValGlyArgGlyAlaPheGlyValValCysLysAlaLysTrpArgAlaLysAspVal 60
 |||||
 Db 283 GTTGTGTGAAGAGAGCCTTTGGAGTTGTTTGAAGCTTAAGTGGAGAGCAAAAGATGTT 342
 |||||
 QY 61 AlaIleLysGlnIleGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
 |||||
 Db 343 GCTATTAAACAAATAGAAAGTGAATCTGAGAGAAAGCGTTTATGTAGAGCTTCGCGAG 402
 |||||
 QY 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnProVal 100
 |||||
 Db 403 TTATCCCGTGTGAACCATCTCTAATATTGTAAGCTTTATGGAGCTTCTTGAATCCAGTG 462
 |||||
 QY 101 CysLeuValMetGluTyrAlaGluGlyCysSerLeuTyrAsnValLeuHisGlyAlaGlu 120
 |||||
 Db 463 TGTCTTGTGATGAATATGCTGAAGGGGCTCTTTATATAATGTGCTGCATGCTGCTGAA 522
 |||||
 QY 121 ProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeuGlnCysSerGlnGly 140
 |||||
 Db 523 CCAITGGCATTATTACTGCTGCCCGCAATGAGTGGTGTGTGTGTGTGTGTGTGTGTGTGT 582
 |||||
 QY 141 ValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysProPro 160
 |||||
 Db 583 GTGGCTTATCTTCACAGCATGCAACCCAAAGCGCTAATTCACAGGACCTGAAACACCA 642
 |||||
 QY 161 AsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCys 180
 |||||
 Db 643 AACTTACTGCTGTTGAGGGGGGACAGTTCTAAAAATTTGTGATTTTGTGTACAGCCTGT 702
 |||||
 QY 181 AspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGluVal 200
 |||||
 Db 703 GACATTTCAGACACATGACCAATAACAAAGGGAGTGTGCTTGGATGGCACCTGAAAGTT 762
 |||||
 QY 201 PheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrpGlyIleIleLeuTrp 220
 |||||
 Db 763 TTTGAAGTAGTAATTACAGTGAATAATGACGCTCTTCAGCTGGGTATTATTCTTTGG 822
 |||||
 QY 221 GluValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMet 240
 |||||
 Db 823 GAAGTGATAACGCGTCGGAAACCCCTTTGATGAGATTGGTGGCCCGCTTTCCGAATCATG 882
 |||||
 QY 241 TrpAlaValHisAsnGlyThrArgProLeuIleLysAsnLeuProLysProIleGlu 260
 |||||
 Db 883 TGGGCTGTTCATAATGTTGTTACTCGACCCACCTGATGATAAAAAATTTACCTAAGCCCATTTGAG 942
 |||||
 QY 261 SerLeuMetThrArgCysTrpSerLysAspProSerGlnArgProSerMetGluGluIle 280
 |||||
 Db 943 AGCCTGATGACTCGTTGTTGTCTAAGATCTCTCCAGCGCCCTTCAATGGAGAAATT 1002
 |||||
 QY 281 ValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyr 300
 |||||
 Db 1003 GTGAAAAATAATGACTCACTTGTGCGGTACTTTCCAGGAGCAGATGAGCCATTACAGTAT 1062
 |||||
 QY 301 ProCysGlnTyrSerAspGluGlnSerAsnSerAlaThrSerThrGlySerPheMet 320
 |||||
 Db 1063 CTTGTCTAGTATTGAGATGAAAGGACAGAGCAACTCTGCCACAGTACAGGCTCATTTATG 1122
 |||||
 QY 321 AspIleAlaSerThrAsnThrSerAsnLysSerAspThrAsnMetGluGlnValProAla 340
 |||||
 Db 1123 GACATTGCTTCTACAAATACGAGTAACAAAGTGAACATAATATGGAGCAAGTTCTTGCC 1182
 |||||
 QY 341 ThrAsnAspThrIleLysArgLeuGluSerLysLeuLeuLysAsnGlnAlaLysGlnGln 360
 |||||


```
Qy 61 AlaIleIysGlnIleGluSerGluSerGluArgIysAlaPheIleValGluIleuArgGln 80
Db 328 GCTATTAAACAGATAGAAAGTGAGTCTGAGAGGAAGGCTTTCAATTGTGAGGCTCGGCGAG 387
Qy 81 LeuSerArgValAsnHisProAsnIleValIysLeuTyrGlyAlaCysLeuAsnProVal 100
Db 388 TTGTGGGTGTGAACCATCTACATTTGTCAAGTTGTACGGAGCCTGCCTGAATCCAGTA 447
Qy 101 CysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGlyAlaGlu 120
Db 448 TGTCTGTGTGTAATATGACAGAGGGGGCTCAATTGTATATATGCTGTGATGTTGCTGAA 507
Qy 121 ProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeuGlnCysSerGlnGly 140
Db 508 CCATTGCTTACTACATGCTGCTCATGTCATGAGCTGTGTGTACAGTTGTTCCTCCAAAGA 567
Qy 141 ValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysProPro 160
Db 568 GTGGCTTACCTGCACAGCATGACGCCAAAGCGCTGATTACAGGGACCTCAAGCCCTCCA 627
Qy 161 AsnLeuLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCys 180
Db 628 AACTGTGCTGGTGTGAGGAGGACAGTTCTAAAAATCTGCCGATTTTGGTACAGCTTGT 687
Qy 181 AspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyrMetAlaProGluVal 200
Db 688 GACATCCAAAACACATATGACCAATAATATAGGGAGTCTGCTTGGATGGCGCTGAAGTG 747
Qy 201 PheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrpGlyIleIleLeuTyr 220
Db 748 TTTGAAGGTAGCAATTTACAGTGAAGAGTGTGATGCTTTCAGCTGGGGTATTTATCCTCTGG 807
Qy 221 GluValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMet 240
Db 808 GAAGTGAATAACCGCCGAAACCCCTTCATGAGATCCGTGGCCAGCTTTCAGAAATCATG 867
Qy 241 TrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGlu 260
Db 868 TGGGCTGTTCATATGTCACCTCGACCCACTGATCAAAATAATTTACCTAAGCCCATTTAG 927
Qy 261 SerLeuMetThrArgCysTrpSerLysAspProSerGlnArgProSerMetGluGluIle 280
Db 928 AGCTTGATGACACGCTGTGTGCTAAGACCCATCTCAGCGCCCTTCAATGGAGGAAT 987
Qy 281 ValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyr 300
Db 988 GTGAAAATAATGACTCACTTCATGCGGTACTTCCAGAGCGGATGAGCCATTACAGTAT 1047
Qy 301 ProCysGlnTyrSerAspGluGlyGlnSerAsnSerAlaThrSerThrGlySerPheMet 320
Db 1048 CTTTGTCTAGTACTCTGATGAAGGGCAGAGCAACTCAGCCACCCAGCACAGGCTCGTTTCATG 1107
Qy 321 AspIleAlaSerThrAsnThrSerAsnLysSerAspThrAsnMetGluGlnValProAla 340
Db 1108 GACATTGCTTCTACAAATACCAATATATAAAGTGACACAATATGGACAGGTTCTCTGCC 1167
Qy 341 ThrAsnAspThrIleLysArgLeuGluSerLysLeuLeuLysAsnGlnAlaLysGlnGln 360
Db 1168 ACAACGACACTATTAAACGCTTGAGTCAAAACTGTGTAAAAACCCAGCGCAAGCAACAG 1227
Qy 361 SerGluSerGlyArgLeuSerLeuGlyAlaSerHisGlySerSerValGluSerLeuPro 380
Db 1228 AGTGAATCTGGACGCTGAGCTTGGAGGCTCTCTGTGGAGCAGTGTGGAGAGCTTGCCCC 1287
Qy 381 ProThrSerGluGlyLysArgMetSerAlaAspMetSerGluIleGluAlaArgIleAla 400
Db 1288 CCCACTTCGAGGGCAAGAGGATGATGCTGCATGCTGAAATAGAAGCCAGGATCGTG 1347
Qy 401 AlaThr----- 402
Db 1348 GCGACTGCGACCTATTCCAAGCCTAAACGGGGCCACCGTAAAAACCGCTTCATTTTGGCAAC 1407
Qy 403 -----ThrGlyAsnGlyGlnProArgArgArgSerIle 413
```

```
Db 1408 ATTCTGATGTCCTCTGAGATCGTCATATCAGGTAAACGGCAACCAAGCGGTAGATCCATC 1467
Qy 414 GlnAspLeuThrValThrGlyThrGluProGlyGlnValSerSerArgSerSerSerPro 433
Db 1468 CAAGACTTGACTGTTACTTGGGACAGAACTGTGTCAAGTCAAGTCAGCGGTATCCAGCCCT 1527
Qy 434 SerValArgMetIleThrThrSerGlyProThrSerGluLysProThrArgSerHisPro 453
Db 1528 AGTGTCAAGATGATCACTACCTCAGGACCAACCTCAGAAAGCCAGCTCGCAGTCAACCA 1587
Qy 454 TrpThrProAspAspSerThrAspThrAsnGlySerAspAsnSerIleProMetAlaTyr 473
Db 1588 TGGACCCCTCATGATTCACAGACACCAATGGCTCAGATAACTCCATCCCAATGCGGTAT 1647
Qy 474 LeuThrLeuAspHisGlnLeuGlnProLeuAlaProCysProAsnSerLysGluSerMet 493
Db 1648 CTTACACTGGATCACCAGCTTACAGCCTCTAGCGCGCTGCCAAACTCCAAAGAAATCCATG 1707
Qy 494 AlaValPheGluGlnHisCysLysMetAlaGlnGluTyrMetLysValGlnThrGluIle 513
Db 1708 GCAGTGTTCGAAACAGCACTGTAAAAATGGCACAGGAGTATATGAAGTTCAAACCGAAATC 1767
Qy 514 AlaLeuLeuLeuGlnArgLysGlnLeuValAlaGluLeuAspGlnAspGluLysAsp 533
Db 1768 GCATTGTACTACAGAGAAGCAAGAACTAGTTGCCAGAAATTGGACCGGATGAAAAGGAC 1827
Qy 534 GlnGlnAsnThrSerArgLeuValGlnGluHisLysLysLeuLeuAspGluAsnLysSer 553
Db 1828 CAGCAAAATACATCTCTGCTGTGGTACAGGAACATAAAAGAGCTTTTAGATGAAAACAAAGC 1887
Qy 554 LeuSerThrTyrTyrGlnGlnCysLysLysGlnLeuGluValIleArgSerGlnGln 573
Db 1888 CTTTCTACTTATTACAGCAATGCAAAAAACAACATAGAGTTCATCAGAAAGCCCAACAGCAG 1947
Qy 574 LysArgGlnGlyThrSer 579
Db 1948 AAACGACAGGCACTTCA 1965

RESULT 9
AF218074 1745 bp mRNA linear PRI 18-DEC-2000
LOCUS Homo sapiens TGF beta-activated kinase splice variant d (TAK1)
DEFINITION mRNA, complete cds.
ACCESSION AF218074
VERSION AF218074.1 GI:6746614
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1745)
AUTHORS Dempsey,C.E., Sakurai,H., Sugita,T. and Guesdon,F.
TITLE Alternative splicing and gene structure of the transforming growth factor beta-activated kinase 1
JOURNAL Biochim. Biophys. Acta 1517 (1), 46-52 (2000)
MEDLINE 20568765
PUBMED 11118615
REFERENCE 2 (bases 1 to 1745)
AUTHORS Dempsey,C.E. and Guesdon,F.
TITLE Direct Submission
JOURNAL Submitted (21-DEC-1999) Division of Molecular and Genetic Medicine, University of Sheffield, School of Medicine, Glossop Road, Sheffield S10 2UF, United Kingdom
FEATURES
source 1. .1745
/organism="Homo sapiens"
/db xref="taxon:9606"
gene 1. .1745
/genes="TAK1"
CDS 7. .1482
/genes="TAK1"
/notes="TAK1d"
```

```
/codon_start=1
/product="TGF beta-activated kinase splice variant d"
/protein_id="AAP27652.1"
/db_xref="gi:6746615"
/translation="MSTASASSSSSSAGEMTEAPSOVLNPFEDIDKEIEVEVNGR
GAFVVCCKAKRKAKVAIKQIESESEKAFIVELRQLSRNHPHYKLYACACINPCL
VMEYAEGLSYNVLHGAEPFYTAIAHMSWLCQSGVAYLHSMOKALIHDLKP
NLLVAGTFLVKICDFGTACDIOHMTNNKSAAMWAPVEFESNTSEKDVSWGII
LMEVITRRKPFDEIGPAPRIMAVHNGTRPPLIKNLPKIESLMTKMSKDSQBS
MEJIKMTHTLMRYEPGADPELOYPCOYDEGOSNASTGSEMDIASTVSNKSDTN
MEQVPATNDTIRLESKLIKNOAKOSESGBLSIGASRGSVESLPTSGKMSADM
SEIEKRIATTGNGQPRRRIQDITVGTGTGPGQYSSSSPSVMTITTSPTSEKPTR
SHPWTPDSDTNGSDNSIFMAVLTLDHQLQARTSGTGE"
```

BASE COUNT 526 a 375 c 412 g 432 t

ORIGIN

Alignment Scores:

Pred. No.:	7.79e-162	Length:	1745
Score:	2756.50	Matches:	538
Percent Similarity:	92.93%	Conservative:	1
Best Local Similarity:	92.76%	Mismatches:	1
Query Match:	91.46%	Indels:	40
DB:	9	Gaps:	1

US-09-830-144-2 (1-579) x AF218074 (1-1745)

```
QY 1 MetSerThrAlaSerAlaAlaSerSerSerSerSerSerAlaGluMetIleGlu 20
Db 7 ATGTCTACAGCCTCTGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCGA 66
QY 21 AlaProSerGlnValLeuAsnPhelGluIleAspTyrLysGluIleGluValGluGlu 40
Db 67 GCCCTTCCAGGTCCTCAACTTGAAGAGATCGACTACAGAGAGATCAGAGTGAAGAG 126
QY 41 ValValGluArgGluValAlaPhelGluValAlaGluValAlaGluValAlaGluVal 60
Db 127 GTTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 186
QY 61 AlaIleLysGlnIleGluSerGluSerGluSerGluSerGluSerGluSerGluSerGlu 80
Db 187 GCATTTAACAATAAAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 246
QY 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGluValAlaGluLeuAsnProVal 100
Db 247 TTATCCCGGTGAAGCACTTAATGTAAGCTTTAAGAGAGAGAGAGAGAGAGAGAGAG 306
QY 101 CysLeuValMetGluTyrValAlaGluIleGluSerLeuTyrAsnValLeuHisGluValAlaGlu 120
Db 307 TGTCTGTGATGGAATATGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 366
QY 121 ProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeuGlnCysSerGlnGly 140
Db 367 CCATTGCCAATATATATCTGCTGCCAGCAATAGTTGGGTGTTACAGTGTCCCAAGGA 426
QY 141 ValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysProPro 160
Db 427 GTGGCTTATCTTCACAGCATGCAACCAAGCCCTTAATTCACAGGAGAGAGAGAGAG 486
QY 161 AsnLeuLeuLeuValAlaGluIleGluTyrValLeuLysIleCysAspPhelGluTyrAlaCys 180
Db 487 AACTTACTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 546
QY 181 AspIleGlnThrHisMetThrAsnAsnLysGluSerAlaAlaTyrMetAlaProGluVal 200
Db 547 GACATTTCAGACACACATGACCAATACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 606
QY 201 PheGluLysSerAsnTyrSerGluLysCysAspValPheSerTyrGluIleIleLeuTyr 220
Db 607 TTGGAAGGTAGTAAATACAGTGAAGAAATGAGAGCTTTCAGCTGGAGTATATCTTTGG 666
QY 221 GluValIleThrArgAlaGluProPheAspGluIleGluValProAlaPheArgIleMet 240
Db 667 GAAGTGAATACGCGTGGAAACCTTTGATGAGATTGTGGCCACGTTTCCAAATCAAG 726
```

```
QY 241 TTPAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGlu 260
Db 727 TGGGCTGTTCAATATATGATCTGCACCAACGATTAATAAAATTTACTTAAGCCATTGAG 786
QY 261 SerLeuMetThrArgCysTyrPheSerLysAspProSerGlnArgProSerMetGluGluIle 280
Db 787 AGCTTATGATCTGTTGTTGTTGTTAAAGATCTCTCCAGAGCCCTTCAATGAGAGAAAT 846
QY 281 ValLysIleMetThrHisLeuMetArgTyrPheProGluValAspGluProLeuGlnTyr 300
Db 847 GTGAAATATATGACTCATTGATGCGGTACTTCCAGAGCAGATGAGCATTACAGTAT 906
QY 301 ProCysGlnTyrSerAspGluGluIleGlnSerAsnSerAlaThrSerThrGlySerPheMet 320
Db 907 CTTGTCAATTAATCAATGAAGACAGACCACTGCGCAACAGTACAGCTCATTCATG 966
QY 321 AspIleAlaSerThrAsnThrSerAsnLysSerAspThrAsnMetGluGluValProAla 340
Db 967 GACATTGCTTCTTAATATACAGATTAACAAAGTACACTAATATGAGAGCAAGTCTGCGC 1026
QY 341 ThrAsnAspThrIleLysArgLeuGluSerLysLeuLeuLysAsnGlnAlaLysGlnGln 360
Db 1027 ACAATGATACTATTAAAGCGCTTGAATCAAAATGTTGAAAAATCAGGCAAGCAACAG 1086
QY 361 SerGluSerGluArgLeuSerLeuGluValAlaSerHisGlySerSerValGluSerLeuPro 380
Db 1087 AGTGAATCTGACGCTTAAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1146
QY 381 ProThrSerGluGluLysArgMetSerAlaAspMetSerGluIleGluValAlaArgIleAla 400
Db 1147 CCAACCTCTAGAGGCAAGAGATGATGCTGACATCTGAAATTAACCTAGAGATGAGC 1206
QY 401 AlaThrThrGlyAsnGlyGlnProArgArgArgSerIleGlnAspLeuThrValThrGly 420
Db 1207 GCAACCAACAGGCAACGAGCAGCAAGAGCTGATCCATCCATCCAACTGACTGACTGAT 1266
QY 421 ThrGluProGluGlnValSerSerArgSerSerSerSerSerValArgMetIleThrThr 440
Db 1267 ACAGAACCTGTGAGGAGACAGAGATGATCCAGTCCAGTCCAGTCCAGTCCAGTCCAG 1326
QY 441 SerGlyProThrSerGluLysProThrArgSerHisProTyrThrProAspAspSerThr 460
Db 1327 TCAGGACCAACCTAGAGAAAGCCAACTCGAATCATTCATCCAGCCCTGATGATTCACA 1386
QY 461 AspThrAsnGlySerAspAsnSerIleProMetAlaTyrLeuThrLeuAspHisGlnLeu 480
Db 1387 GATACCAATGATCAGATTAATCTCCATCCAGTGGCTTATCTTACCTGATCCCAACTA 1446
QY 481 GlnProLeuAlaProCysProAsnSerLysGluSerMetAlaValPheGluGlnHisCys 500
Db 1446 ----- 1446
QY 501 LysMetAlaGlnGluTyrMetLysValGlnThrGluIleAlaLeuLeuGlnArgLys 520
Db 1447 -----CAG 1449
QY 521 -GlnGluLeuValAlaGluLeuAspGlnAspGluLysAspGlnGlnAsnThrSerArgLe 540
Db 1450 GCAAGAACTAGTTGCAGAACTGAGCAGCAGATGAAAGAGACAGCAAAATATCATCTGCT 1509
QY 540 ValAlaGlnLysIleLysLeuLeuAspGluAsnLysSerLeuSerThrTyrTyrGlnGln 560
Db 1510 GGTACAGGAACATATAAAGCTTTAGATGAAACAAAGAGCCCTTTCTACTTACACAGCA 1569
QY 560 nCysLysLysGlnLeuGluValAlaArgSerGlnGlnGlnLysArgGlnGlyThrSer 579
Db 1570 ATGCAGAAAACACTAGAGATCATCAGAGTCAAGACAGCAAGAGCAAGCACTTCA 1627
RESULT 10
E38399 E38399 1704 bp DNA linear PAT 31-JAN-2002
LOCUS NF-kappa B activation inhibitory drug targeting TAK1 and method for
DEFINITION
```


Db	781	AGCCTGATGACTCGTGTGTGTTCTAAAGATCCTTCCAGCGCCCTTCAATGAGGAAATT	840
Qy	281	VallysileMetThrHisLeuMetArgTyrPheProGlyAlaaspGluProLeuGlnTyr	300
Db	841	GTGAANAATGACTCACTTGATCGGTACTTTCCAGAGCAGATGAGCCATTACAGTAT	900
Qy	301	ProCysGlnTyrSerAspGluGlnSerAsnSerAlaThrSerThrGlySerPheMet	320
Db	901	CCTTGTGATGATTACAGATGAAGGACAGCAACTCTGCCACCACTACAGGCTCATTCATG	960
Qy	321	AspilealaSerThrAsnThrSerAsnLysSerAspThrAsnMetGluGlnValProAla	340
Db	961	GACATTGCTCTCAAAATACAGATGAACAAAGTACACTAATATGGAGCAAGTTCTCTGCC	1020
Qy	341	ThrAsnAspThrIleLysArgLeuGluSerLysLeuLeuLysAsnGlnAlaLysGlnGln	360
Db	1021	ACAAATGATCTATTAGCGCTTAGATCAAAAATGTTGAAAATCAGGCAAGCAACAG	1080
Qy	361	SerGluSerGlyArgLeuSerLeuGlyAlaSerHisGlySerSerValGluSerLeuPro	380
Db	1081	AGTGAATCTGGACGCTTTAAGCTTGGAGCCTCCCGTGGAGCAGTGTGGAGAGCTTGCCC	1140
Qy	381	ProThrSerGluGlyLysArgMetSerAlaAspMetSerGluIleGluAlaArgIleAla	400
Db	1141	CCAACTCTGAGGCAAGAGATGAGTGTGACATGCTGAAATAGAAAGCTAGGATCGCC	1200
Qy	401	AlaThrThr-----	403
Db	1201	GCAACCAAGCCTATTCCAAGCCTAAAGGGGGCCACCGTAAACTGCTTCATTGGCAAC	1260
Qy	404	-----GlyAsnGlyGlnProArgArgGserIle	413
Db	1261	ATTCTGGATGTCCTGAGATCGTATATCAGGCAACGACAGCAACGCTAGATCCATC	1320
Qy	414	GlnAspLeuThrValThrGlyThrGluProGlyGlnValSerSerArgSerSerPro	433
Db	1321	CAAGACTTGACTGPAACTGGAAACAGAACCTGGTCAGGTGACGATAGGTGCATCGTCCC	1380
Qy	434	SerValArgMetIleThrThrSerGlyProThrSerGluLysProThrArgSerHisPro	453
Db	1381	AGTGTGAGATGATTACTCTCAGGACCAACCTCAGAAAAGCCAACTCGAAGTCATCCA	1440
Qy	454	TrpThrProAspAspSerThrAspThrAsnGlySerAspAsnSerIleProMetAlaTyr	473
Db	1441	TGGACCCCTGATGATTCACAGATACCAATGGATCAGATCAACTCCATCCCAATGGCTTAT	1500
Qy	474	LeuThrLeuAspHisGlnLeuGlnProLeuAlaProCysProAsnSerLysGluSerMet	493
Db	1501	CTTACACTGGATCACCAACTA-----	1521
Qy	494	AlaValPheGluGlnHisCysLysMetAlaGlnGluTyrMetLysValGlnThrGluIle	513
Db	1521	-----	1521
Qy	514	AlaLeuLeuLeuGlnArgLys-GlnGluLeuValAlaGluLeuAspGlnAspGluLysAs	533
Db	1522	-----CAGGCAAGAAGTCTAGTTGCAGAACTGGACACGATGAAAGGA	1563
Qy	533	pGlnGlnAsnThrSerArgLeuValGlnGluHisLysLysLeuLeuAspGluAsnLysSe	553
Db	1564	CCAGCAAAATACATCTCGCTGGTACAGGAACATATAAAGCTTTTAGATGAAAAACAAGG	1623
Qy	553	rLeuSerThrTyrTyrGlnGlnCysLysLysGlnLeuGluValIleArgSerGlnGlnI	573
Db	1624	CCTTTCTACTTACTACAGCAATGCAAAAAACAACTAGAGGTATCAGAAAGTCAGCAGCA	1683
Qy	573	nLysArgGlnGlyThrSer	579
Db	1684	GAAACGACAGGCCTTCA	1702

RESULT 12

AC114407

LOCUS

AC114407 135147 bp DNA linear HTG 26-JUN-2002

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Mus musculus clone RP23-51G1, WORKING DRAFT SEQUENCE, 5 ordered pieces

AC114407.3 GI:21592111

HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 135147)

Birren,B., Linton,L., Nusbaum,C. and Lander,E.

Mus musculus, clone RP23-51G1

Unpublished

2 (bases 1 to 135147)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,

Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,

Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,

Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,

Cook,A., Cooke,P., DeAtellano,K., Dewar,K., Diaz,J.S., Dodge,S.,

Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,

Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,

Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,

Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,

Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C.,

Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,

McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T.,

Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,

Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,

Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,

Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,

Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S.,

Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,

Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,

Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,

Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,

Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (08-MAR-2002) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 135147)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,

Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,

Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,

Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,

Cook,A., Cooke,P., DeAtellano,K., Dewar,K., Diaz,J.S., Dodge,S.,

Faro,S., Ferreira,P., FitzGerald,M., FitzHugh,W., Gage,D.,

Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L.,

Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,

Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,

Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K.,

Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N.,

Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J.,

Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,

Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P.,

O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,

Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S.,

Schuback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N.,

Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S.,

Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J.,

Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,

Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (26-JUN-2002) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jun 26, 2002 this sequence version replaced gi:21536031.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

QY 409 ArgArgSerIleGlnAspLeuThrValThrGlyThrGluProGlyGlnValSerSer 428
 Db 23282 AGCGTAGATCCATTCAAGACTTGACTGTACTGGACAGAACCTGGTTCAGTGCAGAGT 23341
 QY 429 ArgSerSerProSerValArgMetIleThrThrSerGlyProThrSerGluLysPro 448
 Db 23342 CGGTATCATCCAGCCCTAGAGTGGAAATGATCATCTCTCAGACCAACCTCAGAGAACCA 23401
 QY 449 ThrArgSerHisProTriThrProAspAspSerThrAspThrAsnGlySerAspAsnSer 468
 Db 23402 GCTCGAAGTCAACCGTGGAAACCTGATGATCCACAGATACCAATGGCTTAGATAACTCC 23461
 QY 469 IleProMetAlaTyrrLeuThrLeuAspHisGlnLeuGlnProLeuAlaProCysProAsn 488
 Db 23462 ATCCCAATGGCATATCTTACACTGGATCACCAGCTCAGCCCTCTAGCGCCGTGCCCAAC 23521
 QY 489 SerLysGluSerMetAlaValPheGluGlnHisCysLysMetAlaGlnGluTyrrMetLys 508
 Db 23522 TCCAAATTAATTCATGGCAGTGTGTTGAACCAATGTTAAATGGCACAGGAGTATATGAAA 23581
 QY 509 ValGlnThrGluIleAlaLeuLeuGlnArgLysGlnGluLeuValAlaGluLeuAsp 528
 Db 23582 GTTCAACCGAAATCGCATTTGTTACTACAGAGAAAGCAAGAGTAGTTGCAGAAATTGGAC 23641
 QY 529 GlnAspGluLysAspGlnGlnAsnThrSerArgLeuValGlnGluHisLysLysLeuLeu 548
 Db 23642 CAGGATGAAAGGACCAAGCAAAATACATCTCTCTGGTACAGGAACATAAAAGCTTTTA 23701
 QY 549 AspGluAsnLysSerLeuSerThrTyrrGlnGlnCysLysLysGlnLeuGluValIle 568
 Db 23702 GATGAAACAAAGACCTTTCTTACTATTATACAGCAATGCAAAACAACTAGAGGTTATC 23761
 QY 569 ArgSerGlnGlnLysArgGlnGlyThrSer 579
 Db 23762 AGAGCCCAACGACAGAAACGACAGGCACTTCA 23794
 RESULT 13
 XLU92030 2812 bp mRNA linear VRT 18-APR-1998
 LOCUS Xenopus laevis TGF-beta-activated kinase TAKI mRNA, complete cds.
 DEFINITION U92030
 ACCESSION U92030
 VERSION U92030.1 GI:3057035
 KEYWORDS
 SOURCE
 ORGANISM
 Xenopus laevis.
 Xenopus laevis.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 Xenopodinae; Xenopus.
 1 (bases 1 to 2812)
 Shibuya,H., Iwata,H., Masuyama,N., Gotoh,Y., Yamaguchi,K., Irie,K.,
 Matsumoto,K., Nishida,E. and Ueno,N.
 Role of TAKI and T8B1 in BMP signaling in early Xenopus development
 EMBO J. 17 (4), 1019-1028 (1998)
 98130593
 MEDLINE 9463380
 PUBLISHED
 REFERENCE 2 (bases 1 to 2812)
 Shibuya,H.
 Direct Submission
 Submitted (05-MAR-1997) Faculty of Sciences, Hokkaido University,
 Nishi 6-chome, Kita 12, Kita-ku, Sapporo, Hokkaido 060, Japan
 JOURNAL
 FEATURES
 Location/Qualifiers
 1..2812
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 341..2191
 /note="TGF-beta-activated kinase"
 /codon_start=1
 /product="TAKI"
 /protein_id="AAC14008.1"
 /db_xref="GI:3057035"
 /translations="MSATSAEMTETPPVNLNFEIDYKEIEVEEVRGRTGVVCKAKW
 RGDVAIKQIESESRKAFIVELRQLSRVNHPIVKLYGACLNVPCLVMEYAGGSLY

NVLGAEPYPYTAHAHMSWCLQCAQGVAYLHSMKPKALIHRLKPPNLLVAGGTVL
 KICFGTACDIQTHTNNKSGAAWPEVFGSNYSBKCDVFSWGIILWEVITRRKPF
 DEIGGAPRIWAVHNGTRPPLIKNLKPIESLWTRCWSKDPQRPSPMEIVKIMTHL
 MOYEPGADVSLQYPCQYSDGQSNASTGSCIDITSTNTSNKSDIHIEPQDFQASAS
 NMTKRIESKLAQHLKQAKOTGESRULSPSSGVSEUSEIRGPPSTLTGTSEGR
 RDSADMSLEARIISASTAYVKPRGHRKTASFGNILDVPEIITAGNGQRRRSQVDL
 TIVTSSQESRNSRSSSRVMTTTSPTDPKPRGLPWAPDESDTSDNSIIPM
 AYLTLDHQLPLAPCNSKESMAVFEQHKMAQYMKVQTEIALLQKQBLIAELDQ
 DEKQONTSLRVQSHKLLDENKSLSTVYQCKKQLEIVRSQQQKRGQTS"
 BASE COUNT 808 a 512 c 669 g 723 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,75e-154 Length: 2812
 Score: 2638.50 Matches: 519
 Percent Similarity: 87.50% Conservative: 20
 Best Local Similarity: 84.25% Mismatches: 28
 Query Match: 87.54% Indels: 49
 DB: 5 Gaps: 7
 US-09-830-144-2 (1-579) x XLU92030 (1-2812)
 QY 12 SerSerSerAlaGlyGluMetIleGluAlaProSerGlnValLeuAsnPheGluGluIle 31
 Db 344 TCTGCCACCTCTGCTGAATGATTGAGACTCCA---CCGGTCTGTAACCTTTGAGGAGATC 400
 QY 32 AspTyrrLysGluIleGluValGluValValGlyArgGlyAlaPheGlyValValCys 51
 Db 401 GACTACAGGAGATTGAGGTGGAAGAGGTAGTTGGAAGAGGAACCTTTGGTGTCTGTCGC 460
 QY 52 LysAlaLysTrpArgAlaLysAspValAlaIleLysGlnIleGluSerGluSerGluArg 71
 Db 461 AAAGTAAATGCGGAGGAAAGATGTGGCAATCAACAAATTTGAAAGCGAATCTGAAGG 520
 QY 72 LysAlaPheIleValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLys 91
 Db 521 AAAGCCTTTATTGTTGAGCTGCGACAGTTGTCTCGTGTTAACCATCCCAATATTGTCAAG 580
 QY 92 LeuTyrrGlyAlaCysLeuAsnProValCysLeuValMetGluTyrrAlaGluGlySer 111
 Db 581 TTGTATGGAGCCTGCTAAATCCGGTATGTTTAGTGATGGAGATGTCAGAGAGGAGGTTCA 640
 QY 112 LeuTyrrAsnValLeuHisGlyAlaGluProLeuProTyrrTyrrThrAlaAlaHisAlaMet 131
 Db 641 TTGTACAAATGTTTTCATGGAGCTCAACCGCTGCTTACTATACAGCTGCTCATCAATG 700
 QY 132 SerTrpCysLeuGlnCysSerGlnGlyValAlaTyrrLeuHisSerMetGlnProLysAla 151
 Db 701 AGTTGGTGTTTACAATGTGCCAAGAGTTGCATATTTACATAGCATGAAGCCGAAGCT 760
 QY 152 LeuIleHisArgAspLysProProAsnLeuLeuValAlaGlyGlyThrValLeu 171
 Db 761 CTAATTCAGAGGAGCTCAAAACCAAACTTGTGCTGGTAGCTGGAGGACACCGTCTT 820
 QY 172 LysIleCysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGly 191
 Db 821 AAGATTGTGACTTTTGGTACAGCTGTGATATTACAGACTCACATGACTAATAACAAAGA 880
 QY 192 SerAlaLysTrpMetAlaProGluValPheGluGlySerAsnTyrrSerGluLysCysAsp 211
 Db 881 AGTGACGATGATGGCTCCAGAGTCTTGAAGGTAGCAATTTACAGCGAGAAATGTGAC 940
 QY 212 ValPheSerTrpGlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGlu 231
 Db 941 GTATTAGTTGGGAATAATTTCTTTGGGAAGTAATAACCGAAGAAAACCTTTTCATGAA 1000
 QY 232 IleGlyGlyProAlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeu 251
 Db 1001 ATGTGTGTCCAGCGTTCGGTATCATGTGGCTGTTCAATGGTACTTCGCCCCACATTA 1060
 QY 252 IleLysAsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspPro 271
 Db 1061 ATTAAAAATTTGCCTAAGCCTATTGAAAGCTTAATGACTCGCTGCTGCTCCAAAGATCCC 1120

QY 272 SerGlnArgProSerMetGluGluIleValIleMetThrHisLeuMetArgTyrPhe 291
 |||||
 |||||
 Db 1121 CCACAAAGACCTTCATGAGGAGATGTCAGATATGACTCATCTAATGACGATATTT 1180
 QY 292 ProGlyValAspGluProLeuGlnTyrProCysGlnTyrSerAspGluGlyIleSerAsn 311
 |||||
 |||||
 Db 1181 CCTGGAGCGAGTGTCTTCTTACAGATCCTTGTCAGTACTCAGATGAAAGGCGAAAGTAT 1240
 QY 312 SerAlaThrSerThrGlySerPheMetAspIleAlaSerThrAsnThrSerAsnLysSer 331
 |||||
 |||||
 Db 1241 TCTGCCACAAAGTACAGATCTCTGCACTTATATACCTCACTAATCTAGTACCAAGAGT 1300
 QY 332 AspThrAsnMetGlu-----GlnValProAlaThrAsnAspThrIleLysArg 347
 |||||
 |||||
 Db 1301 GACATCCATATAGAACCTGGTATTTCCAAAGCAAGTCGAGTAATGACCAATCAACGCA 1360
 QY 348 LeuGluSerLysLeu-----LeuLysAsnGlnAlaLysGlnGlnSerGluSerGly 364
 |||||
 |||||
 Db 1361 ATTGAATCAAAAGCTTCACACAAACCTGAAAAATCAAGCAAGCAACGGGTGAATCGGGT 1420
 QY 365 ArgLeuSerLeuGlyAlaSerHisGlySerSerValGluSerLeu----- 379
 |||||
 |||||
 Db 1421 CGATTAGATTGCTCCCTCCTCTCGGGGTACGAGTGAAGAGCTGTCTGAATACGAGCT 1480
 QY 380 ---ProPro-----ThrSerGluGlyLysArgMetSerAlaAspMetSerGlu 394
 |||||
 |||||
 Db 1481 AGACCAACCATCTACTCTGGGTATCTGAGGGGAAACGATAGTGCAGATGCTGAA 1540
 QY 395 IleGluAlaArgIleAlaAlaThrThr----- 403
 |||||
 |||||
 Db 1541 TTGAAAGCAAGAAATTTCTGCAAGTACAGCATATGTCAGCCTAAACGAGCGCAACGCTAAA 1600
 QY 404 -----GlyAsnGly 406
 |||||
 |||||
 Db 1601 ACTGCTTCATTTGGCAACATCTCGATGCCAGATCCACAGATCATCTTACACAGAGAAACGGG 1660
 QY 407 GlnProArgArgArgSerIleGlnAspLeuThrValThrGlyThrGluProGlyGlnVal 426
 |||||
 |||||
 Db 1661 CACACAGAGGGCAGATCTGTACAAAGACCTCACTATGTTGAAACAGAGTCCAGTCAGAG 1720
 QY 427 -----SerSerArgSerSerSerProSerValArgMetIleThrThrSerGlyPro 443
 |||||
 |||||
 Db 1721 AGCAGAAACGTAGCAGATCGTCTAGTCCAAAGTGAGATGATTACTACCTCAGACACCA 1780
 QY 444 ThrSerGluArgProThrArgSerHisProTyrThrProAspAspSerThrAsn 463
 |||||
 |||||
 Db 1781 ACCCCAGACAAAGCCCCACGCTGACCTTCCATGGGCCCCCGCATGAATCTTCAGATACAAAT 1840
 QY 464 GlySerAspAsnSerIleProMetAlaTyrLeuThrLeuAspHisGlnLeuGlnProLeu 483
 |||||
 |||||
 Db 1841 GGAATCGATTAACCTCACTACTATGGCATACTTACACTGGATCATCAGTTACGCTCTTA 1900
 QY 484 AlaProCysProAsnSerLysGluSerMetAlaValPheGluGlnHisCysLysMetAla 503
 |||||
 |||||
 Db 1901 GCACCGGTCTCAAACTTAAGAATCTATGGCTGTGTTGAACAAACATCGCAAAATGGCA 1960
 QY 504 GlnGluTyrMetLysValGlnThrGluIleAlaLeuLeuGlnArgLysGlnGluLeu 523
 |||||
 |||||
 Db 1961 CAGAGAGTACATGAAGTGCAGACCGAAATGCTTTATTGCTGAAAGAAAGCAAGAGCTT 2020
 QY 524 ValAlaGluLeuAspGlnAspGluLysAspGlnGlnAsnThrSerArgLeuValGlnGlu 543
 |||||
 |||||
 Db 2021 ATTGCAGAACTAGATCAGATGAGTGAAGAAAGACACAGACAAACATCGGCTCGGGTCAAGAA 2080
 QY 544 HisLysLysLeuLeuAspGluAsnLysSerLeuSerThrTyrTyrGlnGlnCysLysLys 563
 |||||
 |||||
 Db 2081 CACAAAAAGCTTTAGATGAAAAATAAAGCTTTCCACATATATACGAAATGTAAAAAA 2140
 QY 564 GlnLeuGluValIleArgSerGlnGlnLysArgGlnGlyThrSer 579
 |||||
 |||||
 Db 2141 CAGTTGAGTCAATCGAAGTCAGACGACGAAACGACAGAAACCTCC 2188

RESULT 14
 HSM800550
 LOCUS
 DEFINITION
 Homo sapiens mRNA; cDNA DKFZp586F0420 (from clone DKFZp586F0420); partial cds.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Clone from S. Wiemann. Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de; sequenced by AGO (Berlin/Germany) within the cDNA sequencing consortium of the German Genome Project.
 This clone (DKFZp586F0420) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at <http://www.mips.biochem.mpg.de/proj/cDNA/>.
 FEATURES
 source
 1.3482
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="DKFZp586F0420"
 /tissue_type="uterus"
 /clone_lib="586 (synonym: hute1). Vector pSport1, host DH10B; sites NotI + SalI/MluI"
 /dev_stage="adult"
 1729..2511
 /gene="DKFZp586F0420"
 1729..2511
 /gene="DKFZp586F0420"
 /note="TGF-beta activated kinase 1a, unspliced"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="CAB43687.2"
 /db_xref="GI:6562171"
 /translation="MDIASTNTSNKSDTNEQVPAITNDTIKRLSKLKNQAKQSES GRSLGASGSVSESLPTSEGRKMSADSEIEARLAATGNGOPRRSIOPLVTGT EPGQVSRSSSPETWMTTSTGSEKPTSHPTPDSTDTGSDNSIPMAVLIDHO LOPLACPNKESMAVFEQHCMAQDEVMVQREIALILQKQELVAELDDQDKQONT SRVQEHKILDDNKSLSITYQCKQKQLEIVINSQOKRGQTS"
 polyA_signal
 polyA_site
 BASE COUNT 1222 a 568 c 620 g 1072 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 3.04e-82 Length: 3482
 Score: 1476.00 Matches: 289
 Percent Similarity: 99.66% Conservative: 0
 Best Local Similarity: 99.66% Mismatches: 1
 Query Match: 48.97% Indels: 0
 DB: 9 Gaps: 0
 US-09-830-144-2 (1-579) x HSM800550 (1-3482)
 QY 290 TyrPheProGlyAlaAspGluProLeuGlnTyrProCysGlnTyrSerAspGluGlyGln 309
 |||||
 |||||
 Db 1639 TACTTTCAGAGGAGAGATGAGCCATTACAGATCTTGTGACGATTTGAGTGAAGGACAG 1698
 QY 310 SerAsnSerAlaThrSerThrGlySerPheMetAspIleAlaSerThrAsnThrSerAsn 329
 |||||
 |||||
 Db 1699 AGCAACTGTCACACAGTACAGGCTCATGACACTTCTTCAATAATACAGATTAAC 1758
 QY 330 LysSerAspThrAsnMetGluGlnValProAlaThrAsnAspThrIleLysArgLeuGlu 349

Db	1759	AAAAAGTGACACTAATATGGAGCAAGTTCCTGCCACAATGATCTATTAAAGCGCTTAGAA	1818	
Qy	350	SerLysLeuLeuLysAsnGlnAlaLysGlnSerGluSerGlyArgLeuSerLeuGly	369	
Db	1819	TCAAAATGTTGAAAAATCAGGCACACAGAGTAATCTGGAGCTTTAAGCTTTGGGA	1878	
Qy	370	AlaSerHisGlySerSerValGluSerLeuProProThrSerGluGlyLysArgMetSer	389	
Db	1879	GCCTCCGTCGGAGCAGTGTGAGAGCTTGCCCCCAACCTCTCAGGGCAAGAGATGAGT	1938	
Qy	390	AlaAspMetSerGluIleGluAlaArgIleAlaAlaThrThrGlyAsnGlyGlnProArg	409	
Db	1939	GCTGACATGCTCTGAAATAGAGCTAGGATCGCGCAACACAGCAACGAGCAAGCA	1998	
Qy	410	ArgArgSerIleGlnAspLeuThrValThrGlyThrGluProGlyGlnValSerSerArg	429	
Db	1999	CGTAGATCCATCCAGACCTTGACTGTAACCTGGACAGAACCTGGTCAGGTGACGATAG	2058	
Qy	430	SerSerSerProSerValArgMetIleThrThrSerGlyProProThrSerGluLysProThr	449	
Db	2059	TCATCAGTCCCACTGTGTGAGATGATTAATCTACCTCAGGACCAACCTCAGAAAAGCC	2118	
Qy	450	ArgSerHisProThrThrProAspAspSerThrAspThrAsnGlySerAspAsnSerIle	469	
Db	2119	CGAAGTCATCCATGGACCCCTGATGATTCACAGATACCAATGGATCAGATCACTCCAT	2178	
Qy	470	ProMetAlaTyrLeuThrLeuAspHisGlnLeuGlnProLeuAlaProCysProAsnSer	489	
Db	2179	CCAATGGCTTATCTTACCTGGATCACCACCTACAGCTTAGCCCGTCCCAACTCC	2238	
Qy	490	LysGluSerMetAlaValPheGluGlnHisCysLysMetAlaGlnGluTyrMetLysVal	509	
Db	2239	AAAGAACTCATGGCAGTGTGTGAACAGCATTTGTAATGGCAAGATATATGAAAGTT	2298	
Qy	510	GlnThrGluIleAlaLeuLeuGlnArgLysGlnGluLeuValAlaGluLeuAspGln	529	
Db	2299	CAACAGAAATTCATGTTTATTACAGAAAGCAAGAACTAGTTCAGAACTGGACCAG	2358	
Qy	530	AspGluLysAspGlnGlnAsnThrSerArgLeuValGlnGlnHisLysLysLeuLeuAsp	549	
Db	2359	GATGAAAGAGCAGCAAAATACATCTCGCTGGTACAGAAACATATAAAGCTTTAGAT	2418	
Qy	550	GluAsnLysSerLeuSerThrTyrGlnGlnCysLysLysGlnLeuGluValIleArg	569	
Db	2419	GAAACAAAGAGCTTTCTACTTACTACCAAGCAATGCAAAAAACAATAGAGTCAACAG	2478	
Qy	570	SerGlnGlnGlnLysArgGlnGlyThrSer	579	
Db	2479	AGTCAGCAGCAGAAACACAGCACTTCA	2508	
RESULT 15				
AF199466				
LOCUS				
DEFINITION				
Drosophila melanogaster TGF-beta activated-kinase 1 homolog mRNA,				
complete cds.				
ACCESSION				
AF199466				
VERSION				
AF199466.1 GI:6289098				
KEYWORDS				
SOURCE				
Drosophila melanogaster.				
ORGANISM				
Drosophila melanogaster				
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;				
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
Ephydroidea; Drosophilidae; Drosophila.				
REFERENCE				
1 (bases 1 to 3349)				
AUTHORS				
Takatsu, Y., Nakamura, M., Stapleton, M., Danos, M. C., Matsumoto, K.,				
O'Connor, M. B., Shibuya, H. and Ueno, N.				
TITLE				
TAK1 participates in c-Jun N-terminal kinase signaling during				
Drosophila development				
JOURNAL				
Mol. Cell. Biol. 20 (9), 3015-3026 (2000)				
MEDLINE				
20221548				
PUBMED				
10757786				
REFERENCE				
2 (bases 1 to 3349)				

AUTHORS				
Takatsu, Y., Nakamura, M., Stapleton, M., Danos, M., Matsumoto, M.,				
O'Connor, M. B., Shibuya, H. and Ueno, N.				
TITLE				
Direct Submission				
JOURNAL				
Submitted (28-OCT-1999) Developmental Biology, National Institute				
for Basic Biology, 38 Nishigonaka, Myodaiji-cho, Okazaki, Aichi				
444-8585, Japan				
FEATURES				
source				
1. 3349				
Location/Qualifiers				
/organism="Drosophila melanogaster"				
/db_xref="taxon:7227"				
/chromosome="X"				
/map="19E5-7"				
916. .2952				
/note="TAK1 homolog"				
/codon_start=1				
/product="TGF-beta activated-kinase 1 homolog"				
/protein_id="AAF06815.1"				
/db_xref="GI:6289099"				
/translation="MATASLDALQAAYVDFSEITLREKVGHSYGVVCAVWRDKLVA				
VKFFFAAEOKDIEKEVKLSRVKHPNIIALHGISSYQOATYLI MEFASGGSLHNFH				
GKVKPAYSLAHAMSWARQCAEGLAYLHMTPKPLIHRDVKPLNLLTNGRNLKICDF				
GTVADKSTNNNRGSAAMWAEVFEQSKTEKCDIFSWAIVLWELSVSRKQPPKPGIDN				
AYTIQWKIYKGERPPLLTTCPKRIEDLMTACWKTPEDRPSMQYIVGVNHEIVKDYTG				
ADKALEYTFVNQIVTKESDGTVAQFDSLSQEGELSPSSQTOLTTTAANANVAIA				
ISKTTSSMTENTSSDITPTNSGOLDNNPLFYVWTRWDAIPEESNESRNDSEFN				
LTSABATORLETIRNGMILMACKPWEQLTLDVEANGFDLSPESSSSSTNKAISDRE				
RLVTDTPVMVMTTDLNNNGGIHAHNSGLLSHANGWAQARDBELQREHEQEIFVNSLD				
VYDVPDEBNDGTQEALIELDPELPFPINDAESQLIYRDRHMAKEYLUSVDTNL				
YAAQDFDKLIVQMDRHEQKELLRMKDEGLQSLNNLQOQYASRLAAGHHHPQ				
PHPHPHNPQLQHPHPPMHFLQDECGILLPGSVCGGSESVBEGWVVIPIPHNA"				
BASE COUNT 934 a 893 c 853 g 669 t				
ORIGIN				
Alignment Scores:				
Pred. No.: 5.15e-51 Length: 3349				
Score: 971.00 Matches: 232				
Percent Similarity: 52.22% Conservative: 97				
Best Local Similarity: 36.83% Mismatches: 202				
Query Match: 32.22% Indels: 100				
DB: 3 Gaps: 14				
US-09-830-144-2 (1-579) x AF199466 (1-3349)				
Qy	17	GluMetIleGluAlaProSerGlnValLeuAsnPheGluGluIleAspTyrLysGluIle	36	
Db	913	GAGATGCCACAGCATCGCTGGACGCACTCCAGGCAGCCTATGTGGACTTCAGTGAGATA	972	
Qy	37	GluValGluGluValValGlyArgGlyAlaPheGlyValValCysLysAlaLysTrpArg	56	
Db	973	ACACTAAGAGAGAAAGTCGGCCATGGGTCTCGTGGTGTGCTGCAAGGCCGCTTTGGCG	1032	
Qy	57	AlaLysAspValAlaIleLysGlnIleGluSerGluSerGluArgLysAlaPheIleVal	76	
Db	1033	GACAGCTGGTTCGGTCAAGGAGTTCCTCCAGCGCCGAGCAGAGGACATCGAGAAG	1092	
Qy	77	GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGly-----	94	
Db	1093	GAGGTGAAGCAGTTGTGCGCGTGAAGCACCACCCAAATATATCGCTCTGCACGGATATCC	1152	
Qy	95	AlaCysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsn	114	
Db	1153	TCGTACCAGCAGGCCACCTACCTGATAATGGATTCCGCGAAGGTGGATCGCTGCACAAC	1212	
Qy	115	ValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCys	134	
Db	1213	TTCTCTCACCGC---AAGGTGAAGCCCGGCATATTCCTGCCCCACCCCATGAGCTGGCG	1269	
Qy	135	LeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHis	154	
Db	1270	CGCAATGTGCAGAGGTCTGGCATATTTGCATGCCATGACGCCAAACCACTAATACAT	1329	
Qy	155	ArgAspLeuLysProProAsnLeuLeuValAlaGlyThrValLeuLysIleCys	174	

Db 1330 CGCGACGTGAAGCCGCTGAACTGCTTGTGACCAACAAAGGACGCAATCTGAAGATATCC 1389
 QY AspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaIle 194
 Db 1390 GACTTCGGCAGCGGTGGCGGACAAAGTCGACCATGATGACCAACAAATCGGGATAGTCCGCT 1449
 QY TrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSer 214
 Db 1450 TGGATGGCGCCCGAGGTCTTCGAAAGCTCCAAAGTATACGGAGAACTGTGACATTTTCAGC 1509
 QY TrpGlyIleIleLeuThrGluValIleThrArgArgLysProPheAspGluIleGly 234
 Db 1510 TGGGCATTGTTCTATGGAGGTCTGTCCAGGAAGCGCCCTTTAAAGCATTCACAAAT 1569
 QY ProAlaPheArgIleMetThrAlaValHisAsnGlyThrArgProProLeuIleLysAsn 254
 Db 1570 ---GCCCTACACCATTCAGTCGGAAGATCTACAAAGGTGAACGCCGCCGCTGTGACCACT 1626
 QY LeuProLysProIleGluSerLeuMetThrArgCysTTPSerLysAspProSerGlnArg 274
 Db 1627 TCCCCCAAGCGCATCGAGGACCTGATGACCGCTGTGAAAAACGGTGCAGGATGCG 1686
 QY ProSerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAla 294
 Db 1687 CGGTGCATGACGTACATGATGGGCGCTTATGACAGAGATGTCAGAGACTATACGGGGCG 1746
 QY AspGluProLeuGlnTyr----- 300
 Db 1747 GACCAAGCCCTGGAATACATTTGTTATCAACAGATTGTACCAAAAGAGACGACGCG 1806
 QY 301 -----ProCysGlnTyrSerAspGluGlyGlnSerAsn----- 311
 Db 1807 ACGTGGCCGCTCACCAGGATAGCTCAGTTCGACAGAGGGGAACTGAGCCCTCGTCC 1866
 QY 312 -----SerAla 313
 Db 1867 ACACAGTTAAACCGCAACGCGGCCCAACGCCAATGTGAACGGATAGCAATATCAAAA 1926
 QY ThrSerThrGlySerPheMetAspIleAlaSerThrAsnThrSerAsnLysSerAspThr 333
 Db 1927 ACACAGACTAGCTCATGACCGAAATACCTCATCAACATCATCGGACATCACGCCGACG 1986
 QY AsnMetGluGlnValPro-----AlaThrAsn-----AspThr 344
 Db 1987 AACTCGGGCCACTGACCAATATCCGCTATTCTATATGTCACCAATGCTGGGACGCG 2046
 QY IleLysArgLeuGluSerLysLeuLeuLysAsn----- 355
 Db 2047 ATTCCCGAGAGAGAGCAACGAGAGCCGGAACGATAGCTTCAACCTCACTCGTCCGCT 2106
 QY GlnAlaLysGlnInserGluSerGlyArgLeuSerLeuGlyAlaSerHisGlySerSer 375
 Db 2107 GAGGCCACTCAGCGCTCGGAAACGATCCGGAACGCGCATGCTCTGATGCTGCAAGCCC 2166
 QY ValGluSerLeuProProThrSerGluGlyLysArgMetSerAlaAspMetSerGluIle 395
 Db 2167 ATGGAGCAGCTC-----ACCCTCGACGTGGAGCGCAATGGCTTGTGATCTGAGTCCACG 2220
 QY GluAlaArgIleAlaIleAlaThrThrGlyAsnGlyGlnProArgArgSerIleGlnAsp 415
 Db 2221 GAAAGCGACGACGACGACCAACGCAAGAGCATGGCCGCAACGA----- 2268
 QY LeuThrValThrGlyThrGluPro----- 423
 Db 2269 CTCACGCTGACGACCAACGACCGGTGATGATGACCAAGCATCTGTCCAACAACAGCGC 2328
 QY Gly-GlnValSerSerSerSerSerProSerValArgMetIleThrThrSerGlyPro 443
 Db 2329 GGCATCCACCGCCCACTGAAACGACTGCTGAGCCATGCGAATGTT-----GGCAA 2379
 QY cThrSerGluLysProThrArgSerHisPro-----TrpTh 455
 Db 2380 GCAAGAGATGAGAGCTGCGAGAGCAAGACATGAGCAGAGATTTGTCAACTCGTTTGAC 2439

QY 455 rProAspAspSerThrAspThrAsnGlySerAspAsnSerIleProMetAlaTyrLeuTh 475
 Db 2440 GTC-GACGTGGATCCGACGAGGATGAGAAAGACGACCGACCAAGTCACTGCCGAGAT 2498
 QY rLeuAspHisGlnLeuGlnProLeuAlaProCysProAsnSerLysGluSerMetAlaVal 495
 Db 2499 TCTTGATCCGGAAGCTCCAGCCAGAGCCCGCATACCCAAACGATGCCGAATGCCAGCTCAT 2558
 QY 495 lPheGluGlnHisCysLysMetAlaGlnIleTyrMetLysValGlnThrGluIleAlaLe 515
 Db 2559 CTACCGGACACACGACACATGGCCCAAGGATACCTGAGCGTGCACACGACCTTACTTA 2618
 QY 515 uLeuLeuGlnArgLysGlnGluLeuValAlaGluLeuAspGlnAspGluLysAspGlnG 535
 Db 2619 CGCGCAGGACTTCMAAGACAGCTCATCTGCAATGAGACCGAAACGAGCGCAACAGAA 2678
 QY 535 nAsnThrSerArgLeuValGlnGlnHisLysLysLeuLeuAspGluAsnLysSerLeuSe 555
 Db 2679 G-----CAGGAGCTTCTCGCGCAAGATGAAGAGCAAGAGAGGTCTTCA 2720
 QY 555 rThrTyrGlnGlnCysLysLysGln 564
 Db 2721 GAGCTTTACAAACATCTGCAGCAGCAG 2748

Search completed: December 10, 2002, 01:08:33
 Job time : 3724.32 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 9, 2002, 22:57:35 ; Search time 286.56 Seconds
(without alignments)
4550.213 Million cell updates/sec

Title: US-09-830-144-2

Perfect score: 3014

Sequence: 1 MSTASASSSSSSAGEMIE.....QCKQLVIRSQOQKRGTS 579

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlp
-O=/cgn2_1/USPTO_spool/US09830144/runat_04122002_141353_2242/app_query.fasta_1.1422
-DB=N_Geneseq_101002 -QFMT=fastac -SURFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09830144.ccgcn_1_113 @runat_04122002_141353_2242 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MWAP -LARGQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_101002.*

1: /SID32/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID32/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID32/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID32/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID32/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID32/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID32/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID32/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID32/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID32/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID32/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID32/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID32/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID32/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID32/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID32/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID32/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID32/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID32/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID32/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID32/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID32/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID32/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID32/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match %	Score	Length	DB	ID	Description
1	3014	100.0	1788	20	AA56285	Human TAK1-6xHis e
2	3014	100.0	1959	18	AAT85095	Human transforming
3	3014	100.0	2656	20	AA56279	Human TAK1 encodin
4	3014	100.0	2656	21	AAA39105	Human TAK-1 nucleo
5	3006	99.7	2769	24	ABL88437	Pain regulated cDN
6	3006	99.7	2785	20	AA59696	Human TGF-beta act
7	2986	99.1	2443	18	AAT85094	Mouse transforming
8	2982.5	99.0	2866	20	AA59697	Human TGF-beta act
9	2746	91.1	1704	20	AA59698	Human TGF-beta act
10	971	32.2	3367	23	ABL02489	Drosophila melanog
11	702.5	23.3	10997	23	ABL02488	Drosophila melanog
12	505	16.8	759	23	ABL08337	Drosophila melanog
13	493.5	16.4	3454	24	ABL70018	Pancreas cancer re
14	490	16.3	2120	21	AA599726	cDNA encoding huma
15	482	16.0	1706	21	AA5674	DNA encoding a hum
16	482	16.0	2191	22	AAH99263	Human protein enco
17	482	16.0	2194	21	AA599734	Cardiovascular sys
18	482	16.0	2220	21	AA593783	Human survival reg
19	481	16.0	1365	21	AA599727	cDNA encoding huma
20	481	16.0	2272	21	AA599736	Cardiovascular sys
21	469	15.6	3141	22	AAD18824	Human kinase (PKIN
22	469	15.6	3538	24	AAD34309	Human PKIN-12 cDNA
23	462	15.3	3066	24	ABO86165	Novel human gene.
24	462	15.3	3558	24	ABK83874	Human cDNA differe
25	462	15.3	3558	24	AAH46139	Human mitogen acti
26	461.5	15.3	2157	22	AAH46913	cDNA encoding huma
27	461.5	15.3	3111	24	ABN86357	Novel human protei
28	461.5	15.3	3518	24	ABN86358	Novel human protei
29	457.5	15.2	2622	22	AAF75336	Human TGF-beta rec
30	457.5	15.2	3967	22	AAH73366	Human cervical can
31	456.5	15.1	2403	22	AAF44701	Novel protein kina
32	436.5	14.5	3072	23	ABL04365	Drosophila melanog
33	434.5	14.4	1224	21	AAC43254	Arabidopsis thalia
34	434	14.4	1063	22	AAH34976	Human colon cancer
35	434	14.4	1631	21	AAC39537	Arabidopsis thalia
36	427.5	14.2	1591	21	AAC40839	Arabidopsis thalia
37	426.5	14.2	2283	21	AAC48526	Arabidopsis thalia
38	425	14.1	2254	21	AA599735	Cardiovascular sys
39	424	14.1	3109	23	ABL29755	Drosophila melanog
40	418.5	13.9	3025	21	AAA47606	Human CARK (Cardia
41	418	13.9	2069	21	AA599737	Cardiovascular sys
42	418	13.9	2505	21	AAA47607	Human CARK (Cardia
43	418	13.9	2508	22	AAF44702	Novel protein kina
44	418	13.9	2788	23	AA580720	DNA encoding novel
45	418	13.9	3026	21	AAA47608	Rat CARK (Cardiac

ALIGNMENTS

RESULT 1
AA56285
ID AA56285 standard; DNA; 1788 BP.
XX
AC AA56285;
XX
DT 21-JUL-1999 (first entry)
XX
DE Human TAK1-6xHis encoding DNA.
XX
KW Human; TAB1; TAK1; screening; inhibition; TGF-beta;
KW transforming growth factor beta; ss.
XX
OS Homo sapiens.
XX
FT Synthetic.
XX
Key Location/Qualifiers
CDS 7..1779

```

FT      /cag= a
XX      MO9921010-A1.
XX      29-APR-1999.
XX      22-OCT-1998; 98WO-JP04796.
XX      22-OCT-1997; 97JP-0290188.
XX      (CHUS ) CHUGAI SEIYAKU KK.
XX      Ohtomo T, Ono K, Tsuchiya M;
XX      WPI; 1999-312645/26.
XX      P-PSDB; AAY09547.
PT      Screening for TGF- beta inhibitory substances, which are useful as
PT      drugs for treatment of diseases relating to its disorder
XX      Example 1; Page 167-171; 195pp; Japanese.
XX      A method has been developed for screening for substances which inhibit
XX      the binding of TAK1 polypeptide to TAB1 polypeptide. The method
XX      comprises: (a) contacting the polypeptide in the presence of a sample;
XX      and (b) detecting the amount of bound polypeptide, in which the sample
XX      can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming
XX      growth factor (TGF)-beta inhibitory substances can be used in drugs for
XX      indications e.g. as TGF-beta signal transmission inhibitors or
XX      activators, or extracellular matrix protein production enhancement
XX      inhibitors or activators, or cell proliferation prevention inhibitors or
XX      activators, or monocyte migration inhibitors or activators, or
XX      physiological activity induction inhibitors or activators, or
XX      immunosuppression inhibitors or activators, or amyloid beta protein
XX      precipitation inhibitors or activators, and such substances can also be
XX      inhibitors of the TAK1 polypeptide function, particularly kinase
XX      activity. The present sequence encodes TAK1-6xhis from an example of
XX      the present invention.
XX      Sequence 1788 BP; 551 A; 394 C; 413 G; 430 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 3,296-217 Length: 1788
Score: 3014.00 Matches: 579
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-09-830-144-2 (1-579) x AAY56285 (1-1788)
OY 1 MetSerThrAlaSerAlaIaIaSerSerSerSerSerAlaGlyIuMetIleGlu 20
DB 7 ATGTCACAGCCCTGCGGCTCTCTCTCTCTCTGCGCGGATGATGATCGAA 66
OY 21 AlaProSerGlnValIeuAsnPhenGluGluIleAspTyrLysGluIleGluValGluGlu 40
DB 67 GCCCTTCCAGGCTCTCAACTTGAAGAGATCGACTTAAGAGAGATCGAGGAGAG 126
OY 41 ValValGlyArgGlyAlaPhenGlyValValCysLysAlaLysTyrArgAlaLysAspVal 60
DB 127 GTTGTGGAAAGAGAGCTTTGAGTGTGTTGCAAGAGCTTAAGTGAAGACCAAAAGTGT 186
OY 61 AlaIleLysGlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
DB 187 GCATTAAACAATAGAAAGTGAATCTGAGAGAAAGCCGTTATGTAGAGCTTCGCGAG 246
OY 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnProVal 100
DB 247 TTATCCGCTGTGAACCATCTTAATATGTAAAGCTTTAAGAGAGCCGCTGAATCCAGTG 306
OY 101 CysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnValIeuHisGlyAlaGlu 120

```

```

DB 307 TGTCTTGATGATGAATATGCTGAAGGGGCTCTTATATATATGTCGATGCTGAA 366
OY 121 ProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeuGlnCysSerGlnGly 140
DB 367 CCATTGCCATATATATATGCTGCTCCACGCAATGAGTGGTGTATTACATGTTCCCAAGA 426
OY 141 ValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisAspAspLeuLysProPro 160
DB 427 GTGGCTTATCTTCACAGCATGCAACCAAGGCTTAATTCACAGGAGACTGAAACCAACA 486
OY 161 AsnLeuLeuLeuValAlaGlyIleThrValLeuLysIleCysAspPheGlyThrAlaCys 180
DB 487 AACTTACTGCTGTTGACAGGGGAGACAGTTCTAAATAATTTGATTGTGACAGCTGCT 546
OY 181 AspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyrMetAlaProGluVal 200
DB 547 GACATTCAAGACACATGACCAATTAACAAGGAGATGCTGCTGATGACCTTAAGT 606
OY 201 PheGluGlySerAsnTyrSerGluLysCysAspValPheSerTyrGlyIleIleLeuTyr 220
DB 607 TTGAAGGTATGATTAATTAAGTGAATAATGACGCTTCAGCTGGGGTATTATCTTTGG 666
OY 221 GluValIleThrArgArgLysProPheAspGluIleGlyIleProAlaPheArgIleMet 240
DB 667 GAAGGTAAACGGGTGGAACCCCTTGATGAGATTGGTGCCAGCTTCGAAATCATG 726
OY 241 TyrAlaValHisAsnGlyThrArgProLeuIleLysAsnLeuProLysProIleGlu 260
DB 727 TGGCGTTCATATATGATGCTGACACCACTGATTAATAATTTACTTAAAGCCCATGAG 786
OY 261 SerLeuMetThrArgCysTyrSerLysAspProSerGlnArgProSerMetGluGluIle 280
DB 787 AGCTGATGACTGCTGTTGCTTAAGATCTTCCACAGCGCTTCATGAGGAAAT 846
OY 281 ValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyr 300
DB 847 GTGAAATATATGACTCATGATGCGGTACTTCCAGGAGCATAGCATTAACGATTA 906
OY 301 ProCysGlnTyrSerAspGluGluIleGlnSerAsnSerAlaThrSerThrGlySerPheMet 320
DB 907 CTTGTGATTAATCAAGTGAAGAGAGAGACTTGCACACGTAACAGCTCATTCATG 966
OY 321 AspIleAlaSerThrAsnThrSerAsnLysSerAspThrAsnMetGluGluValProAla 340
DB 967 GACATTCCTCTCAATAATAGAGTAACAAGAGACACTTAATGAGAGCAATCTCTGCC 1026
OY 341 ThrAsnAspThrIleLysArgLeuGluSerLysLeuLeuLysAsnGlnAlaLysGlnGln 360
DB 1027 ACAATGATATCTTAAGCCCTTAAGATCAAAATGTTGAAATATCAGGCAAGCAACAG 1086
OY 361 SerGluSerGlyArgLeuSerLeuGlyAlaSerHisGlySerSerValGluSerLeuPro 380
DB 1087 AGTAATCTGAGAGCTTAATGCTGGAGCTCCCATGAGAGAGCTGAGAGCTTGCCC 1146
OY 381 ProThrSerGluGlyLysArgMetSerAlaAspMetSerGluIleGluAlaArgIleAla 400
DB 1147 CCAACCTCTGAGGCAAGAGAGATGAGTGCATGCTGAATATGAAGCTGAGATGCC 1206
OY 401 AlaThrThrGlyAsnGlyGlnProArgArgSerIleGlnAsnLeuThrValThrGly 420
DB 1207 GCAACCAAGGCAACGAGAGCCAAAGCGTGAATCCATCCAAAGCTGATGATGATCGGA 1266
OY 421 ThrGluProGlyGlnValSerSerArgSerSerSerProSerValArgMetIleThrThr 440
DB 1267 ACAGAACCTGCTCAGGTGAGAGAGTCAATCCAGTCCAGTGCAGAAATGATTACTACC 1326
OY 441 SerGlyProThrSerGluLysProThrArgSerHisProTyrThrProAspAspSerThr 460
DB 1327 TCAGAGCAACCTCAGAAAGCCCAACTGAAATCATCATGAGAGCCCTGATGATTCACACA 1386
OY 461 AspThrAsnGlySerAspAsnSerIleProMetAlaTyrIleuThrLeuAspHisGlnLeu 480
DB 1387 GATACCAATGATGATGATTAATCTCATCCCAATGAGCTTATCTTACATGATCAACCACTA 1446

```


QY	321	AspIleAlaSerThrAsnThrSerAsnLysSerAspThrAsnMetGlnGlnValProIleA	340
Db	1143	GACATTCCTCTCAAAATACAGAGTAACAAAAGGACACCTAATATGGAGCAAGTTCCTCC	1202
QY	341	ThrAsnAspThrIleLysAsArgLeuGluSerLysLeuLeuLysAsnGlnAlaLysGln	360
Db	1203	ACAAATGATACTATTAAAGCCCTTGGAAATCAAAATTTGTGAAAATATCAGCAAAAGCAACG	1262
QY	361	SerGluSerGlyArgLeuSerLeuGlyAlaSerHisGlySerSerValGluSerLeuPro	380
Db	1263	AGTGATCTGGACGTTTAAGCTTGGAGCCTCCCATGGAGAGCAGTGTGAGAGCTTGGCC	1322
QY	381	ProThrSerGluGlyLysArgMetSerAlaAspMetSerGlnIleGlnAlaArgIleAla	400
Db	1323	CCAACCTCTGTAGGGCCAGAGAGATGAGTGACATGTCTGAATATGAAGCTAGAGTCGC	1382
QY	401	AlaThrThrGlyAsnGlyGlnProArgArgSerIleGlnAspLeuThrValThrGly	420
Db	1383	GCAACACAGGCAACGGACACCCAGACGTAGATCCATCCAGACTTGACTTAATCTGA	1442
QY	421	ThrGluProGlyGlnValSerSerArgSerSerSerProSerValArgMetIleThrThr	440
Db	1443	ACAGAACTGTGTCAGGTGAGAGTAGATCATCCAGTCCCGATGTCGAATGTATTCTACC	1502
QY	441	SerGlyProThrSerGluLysProThrArgSerHisProThrProAspSerThr	460
Db	1503	TCAGAGCAACACCTCAGAAAACCCAACTCGAAGTCATCCAGAGACCCCTGATATTCACAA	1562
QY	461	AspThrAsnGlySerAspAsnSerIleProMetAlaTyrLeuThrLeuAspHisGlnLeu	480
Db	1563	GATACCAATGGATAGATTAATCTCATCCCAATGGCTTATCTTACATGGATCACCAACTA	1622
QY	481	GlnProLeuAlaProCysProAsnSerLysGluSerMetAlaValPheGluGlnHisCys	500
Db	1623	CAGCCTTAGACACCGTGGCCCAATCCCAAGAAATCTATGGCAGTGTGTAACAGCAATTGT	1682
QY	501	LysMetAlaGlnGluTyrMetLysValGlnThrGlnIleAlaLeuLeuGlnArgLys	520
Db	1683	AAAAAGGCAACAAGATATATGAAAGTTCAACACGAAATTCATTTGATTATACAGAGAAAG	1742
QY	521	GlnGluLeuValAlaGluLeuAspGlnAspGluLysAspGlnIleAsnThrSerArgLeu	540
Db	1743	CAGAACTAGTTGCCGAACCTGGACACAGATGAAAAGAACACGAAAATATCATCTGGCCG	1802
QY	541	ValGlnGluHisLysLysLeuLeuAspGluAsnLysSerLeuSerThrTyrTyrGlnGln	560
Db	1803	GTACAGGAACATATAAAAGCTTTTAGATGAAAAACAAAGCCCTTCTACTACTCCAGCAA	1862
QY	561	CysLysLysGlnLeuGlnValIleArgSerGlnGlnIleLysArgGlnGlyThrSer	579
Db	1863	TGCAAAAAACAACCTAGAGGTCAATCAGAAAGTCAGCAGCAAGAAACGACAGGCACTTCA	1919
RESULT 3			
AAKS56279			
ID	AAKS56279 standard; DNA; 2656 BP.		
XX	AAKS56279;		
AC	21-JUL-1999 (first entry)		
XX	Human TAK1 encoding DNA.		
XX	Human; TAB1; TAK1; screening; inhibition; TGF-beta;		
KW	transforming growth factor beta; ss.		
XX	Homo sapiens.		
OS	Location/Qualifiers		
XX	Key		
XX	FT CDS		
XX	FT		
XX	/tag= a		
XX	WO9921010-A1.		

XX	29-APR-1999.
PF	22-OCT-1998; 98WO-JP04796.
PR	22-OCT-1997; 97JP-0290188.
XX	(CHUS) CHUGAI SEIYAKU KK.
PI	Ohtomo T., Ono K., Tsuchiya M;
DR	WPI; 1999-312645/26.
P-PSDB;	AAV09542.
PT	Screening for TGF-beta inhibitory substances, which are useful as
PT	drugs for treatment of diseases relating to its disorder
PS	Example 1; Page 150-154; 195pp; Japanese.
XX	
CC	A method has been developed for screening for substances which inhibit
CC	the binding of TAK1 polypeptide to TAB1 polypeptide. The method
CC	comprises: (a) contacting the polypeptide in the presence of a sample;
CC	and (b) detecting the amount or bound polypeptide, in which the sample
CC	can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming
CC	growth factor (TGF)-beta inhibitory substances can be used in drugs for
CC	indications e.g. as TGF-beta signal transmission inhibitors or
CC	activators, or extracellular matrix protein production enhancement
CC	inhibitors or activators, or cell proliferation prevention inhibitors o
CC	activators, or monocyte migration inhibitors or activators, or
CC	physiological activity induction inhibitors or activators, or
CC	immunosuppression inhibitors or activators, or amyloid beta protein
CC	precipitation inhibitors or activators, and such substances can also be
CC	inhibitors of the TAK1 polypeptide function, particularly kinase
CC	activity. The present sequence encodes human TAK1.
XX	
SQ	Sequence 2656 BP; 778 A; 557 C; 620 G; 701 T; 0 other;
Alignment Scores:	
Pred. No.:	5, 47e-217 Length: 2656
Score:	3014.00 Matches: 579
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatches: 0
Query Match:	100.00% Indels: 0
DB:	Gaps: 0
US-09-830-144-2 (1-579) x AA56279 (1-2656)	
QY	1 MetSerThrAlaSerAlaAlaSerSerSerSerSerAlaGlyIleMetIleGlu 20
Db	183 ATGTCTACAGGCTCTGGCGGCTCTTCCTCCTCTGCCTTGCGCGGTGAAGATCGAA 242
QY	21 AlaProSerGlnValLeuAsnProHegiLugluIleAerpyrTrlysGuluIegluValGluGlu 40
Db	243 GCCCTTCCCAAGGCCCTCAACTTTGAAGAAGTCGACTCAAGAGAGATCGAGTGGAAAG 302
QY	41 ValValGlyArgGlyAlaPheGlyValValCysLysAlaLSTSPARGalalysASPval 60
Db	303 GTTGTGGACAAGAGCGCTTTGGAGTGTGTTCCAAAGTAAGTGAAGACAAAAGATT 362
QY	61 AlaIleLysGlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
Db	363 GCCTATAAACAAATAGAAATGAACTGTGAGAGAAAGCGTTTATTGTAGAGCTTCGCAG 422
QY	81 LeuSerArgValAsnHisProAsnIleValLysLeuIryGlyAlaCybleuAsnProval 100
Db	423 TTATCCCGTGTGAACCATCTCTAATATTGTGAAAGCTTTATGAGCGCTGTGGAATCCAGTG 482
QY	101 CysLeuValMetGluTrpAlaGluGlyGlySerLeuTrpAsnValIleuHisGlyAlaGlu 120
Db	483 TGCTTGTGATGGAAATVGTGTGAAGGGGGCTCTTTAATTAATGTGCTGCATGTGCTGAA 542
QY	121 ProLeuProTrpTrpThrAlaAlaHisAlaMetSerTrpCysLeuGlnCysSerGlnGly 140
Db	543 CCATGGCATATTATATCTGTGCGCACGGAAGTAGTGGTGTGTTAAAGATGTGTTCCAAAGA 602

QY 141 ValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysProPro 160
Db 603 GTGGCTTATCTTACAGCATGCAACCCAAAGCGCTAATTACACAGGACCTGAAACCA 662
QY 161 AsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCys 180
Db 663 AACTTACTGTGTGGTTGCAGGGGGACAGTCTTAAATTTGTGATTTTGGTACAGCCTGT 722
QY 181 AspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGluVal 200
Db 723 GACATTGAGACACATGACCAATAACACAGGGAGTGCTGCTTGGATGGCACTGAAGTT 782
QY 201 PheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrpGlyIleIleLeuTrp 220
Db 783 TTTGAAGGTAGTAATTACAGTGAATAATGTGACGTCTTCAGCTGGGGTATTATTCTTTGG 842
QY 221 GluValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMet 240
Db 843 GAAGTGATAACGGGTGCGAACCCTTTTGATGAGATTGGTGGCCACGCTTTCGAATCATG 902
QY 241 TrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGlu 260
Db 903 TGGCTGTTCTAATAGTACTCGACCACCACTGATATAAAATTTACTAAGCCCATGTAG 962
QY 261 SerLeuMetThrArgCysTrpSerLysAspProSerGlnArgProSerMetGluGluIle 280
Db 963 AGCCTGATGACTCGTGTGTTGGTCTAAGATCCCTCCAGCGCCCTTCAATGAGGAAAT 1022
QY 281 ValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyr 300
Db 1023 GTGAAAAATAATGACTCACTTGTATGCGGTACTTTCCAGGAGCAGATGAGCCATTACAGTAT 1082
QY 301 ProCysGlnTyrSerAspGluGlyGlnSerAsnSerAlaThrSerThrGlySerPheMet 320
Db 1083 CCTTGTTCAGTATTCAGATGAGACAGACGCAACTCTGCCACCACTACAGGCTCATTCATG 1142
QY 321 AspIleAlaSerThrAsnThrSerAsnLysSerAspThrAsnMetGluGlnValProAla 340
Db 1143 GACATTGCTTCTCAATAACGAGTAACAAAAGTGACACTAATATGAGCAAGTTCTCTGCC 1202
QY 341 ThrAsnAspThrIleLysArgLeuGluSerLysLeuLeuLysAsnGlnAlaLysGlnGln 360
Db 1203 ACAATGATATACTAATAAGCCCTTAGAATCAAAATTTGTGAAAAATCAGGCAAGCAACAG 1262
QY 361 SerGluSerGlyArgLeuSerLeuGlyAlaSerHisGlySerSerValGluSerLeuPro 380
Db 1263 AGTGAATCTGGAGCTTTAAGCTTTGGAGCCCTCCATGGGAGCAGTGTGGAGACTTTGCC 1322
QY 381 ProThrSerGluGlyLysArgMetSerAlaAspMetSerGluIleGluAlaArgIleAla 400
Db 1323 CCAACCTCTGAGGCAAGAGGATGAGTGCTGACATGTCTGAAATAGAAGCTAGGATCGCC 1382
QY 401 AlaThrThrGlyAsnGlyGlnProArgArgSerIleGlnAspLeuThrValThrGly 420
Db 1383 GCAACACAGGCAACGGACAGCAAGCAGTAGATCCATCCAAGACTTGACTGTAACGTGA 1442
QY 421 ThrGluProGlyGlnValSerSerArgSerSerProSerValArgMetIleThrThr 440
Db 1443 ACAGAACCTGTGTCAGTGACGATAGGTCTATCCAGTCCCGTGTGCAATGATTTACTAC 1502
QY 441 SerGlyProThrSerGluLysProThrArgSerHisProTrpThrProAspAspSerThr 460
Db 1503 TCAGGACCAACCTCAGAAAAAGCAACTCGAAGTCATCCATGGACCCCTGATGATCCACA 1562
QY 461 AspThrAsnGlySerAspAsnSerIleProMetAlaTyrIleuThrLeuAspHisGlnLeu 480
Db 1563 GATACCAATGGATCAGATAACTCCATCCCAATGGCTTATCTTACACTGGATCACCACCTA 1622
QY 481 GlnProLeuAlaProCysProAsnSerLysGluSerMetAlaValPheGluGlnHisCys 500
Db 1623 CAGCCTCTAGCACCGTGGCCCAAACTCCAAAGAAATCTATGGCAGTGTGTTGAACAGCATTTGT 1682

QY 501 LysMetAlaGlnGluTyrMetLysValGlnThrGluIleAlaLeuLeuLeuGlnArgLys 520
Db 1693 AAAATGGCACAGAATATATCAAGATTCAACAGAAATTGCATTGTTATTACAGAGAAAG 1742
QY 521 GlnGluLeuValAlaGluLeuAspGlnAspGluLysAspGlnGlnAsnThrSerArgLeu 540
Db 1743 CAAGAACTAGTTGCAGAACTGGACAGGATGAAAAGGACCAAGCAAAATACATCTCGCCTG 1802
QY 541 ValGlnGluHisLysLysLeuLeuAspGluAsnLysSerLysSerLeuSerThrTyrGlnGln 560
Db 1803 GTACAGGAACATAAAAGCTTTTAGATGAAAAACAAGCCCTTTCTACTTACTACAGCAA 1862
QY 561 CysLysLysGlnLeuGluValIleArgSerGlnGlnGlnLysArgGlnGlyThrSer 579
Db 1863 TGCAAAAACAACCTAGAGGTCTATCAGAACTCAGCAGCAGAAACGACAGGCCTTCA 1919
RESULT 4
AAA39105
ID AAA39105 standard; DNA; 2656 BP.
XX
AC AAA39105;
XX
DT 04-SEP-2000 (first entry)
XX
DE Human TAK-1 nucleotide sequence SEQ ID NO:1.
XX
KW Human; TAK-1; TAB-1; mitogen activated protein kinase; MAPK;
KW screening; signal transduction; inhibition; inflammatory cytokine;
KW IL-1; interleukin 1; TNF; tumour necrosis factor; inflammation;
KW antiinflammatory; suppression; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 183..1922
FT /*tag= a
FT /product= "TAK-1"
XX
WO200023610-A1.
XX
PD 27-APR-2000.
XX
PF 21-OCT-1999; 99WO-JP05817.
XX
PR 21-OCT-1998; 98JP-0299962.
XX
PA (CHUS) CHUGAI SEIYAKU KK.
XX
PI Tsuchiya M, Ohtomo T, Sugamata Y, Matsumoto K;
XX
WPI: 2000-339707/29.
DR P-PSDB; AAY91000.
XX
PT Method for screening inhibitors of TAK1 signal transduction for
PT suppression of inflammatory cytokine production and use as
PT antiinflammatory agents -
XX
PS Example 1; Page 73-80; 100pp; Japanese.
XX
CC The present invention describes a method for screening compounds for
CC inhibition of inflammatory cytokine signal transduction by contacting
CC the sample with TAK1 and its receptor TAB1 and selecting for inhibition
CC of TAK1/TAB1 binding. Also described is a method for screening compounds
CC for inhibition of inflammatory cytokine signal transduction in which the
CC inhibition of TAK1 phosphorylation is selected for; and drug
CC compositions for the treatment of inflammatory disorders containing as
CC active component an inflammatory cytokine signal transduction inhibitor.
CC TAK1 is an essential component of the signalling process which results
CC in release of inflammatory cytokines such as interleukin-1 (IL-1),
CC IL-10, tumour necrosis factor (TNF) and IL-6. The methods can be used
CC for the selection of effective antiinflammatory agents. The present
CC sequence encodes human TAK-1, which is used in the exemplification of the
CC present invention.

Db	1663	AAATGCGCAAGAAATATGTGAAGTTCAAACGAAATTGCCTTCTTTACAGAGAAG	1722
Oy	521	GInGluleuValIaGIuleuAspGlnAspGlnLysAspGlnGlnAsnThrSerArgLeu	540
Db	1723	CAGAACTACTAGTCAGAACGTGACCAGGATGAAAAGACACACAAAAATACAATCTGGCCTG	1782
Oy	541	VaIGlInGHIsylsYstLeuAspGluAnuLYSserLeuSerThrTYTyGInGln	560
Db	1783	GTA CAGGAACATAAAGAGCTTTAGATGA AAAACMAAGCCTTCTACTACTACAGCAA	1842
Oy	561	CysHysSGInleuGlnValIIeaRgSerGInGlnLYsaRgInGlyThrSer	579
Db	1843	TGCAAAAACACACTAGAGGTCATCAGAGTCAGACAGACAAACGACAGCATTCA	1899
RESULT 6			
AAX99696	ID	AAX99696 standard; cDNA to mRNA; 2785 BP.	
XX	XX		
AC	XX	AAX99696;	
DT	18-OCT-1999	(first entry)	
DE			
XX		Human TGF-beta activated kinase (TAK) 1a encoding nucleotide sequence.	
KM	Nuclear factor kappa B; NF-kB; inhibitor; TGF-beta activated kinase 1;		
KM	TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;		
KM	intractable disease; atrophic dermatitis; psoriasis; viral infection;		
KM	endotoxin shock; septicemia; human; hTAK1a; ss.		
OS	Homo sapiens.		
FH	Key	Location/Qualifiers	
FT	CDS	163..1902	
FT		/*tag= a	
FT		/product= "hTAK1a"	
PN	MO9940202-A1.		
XX	12-AUG-1999.		
PD			
PF	02-FEB-1999;	99WC-JP00422.	
PR	30-OCT-1998;	98JP-0309316.	
PR	06-FEB-1998;	98JP-0026003.	
PA	(TANABE SEIYAKU CO.		
PI	Hasegawa K, Kageyama N, Sakurai H, Sugita T;		
DR	WPI; 1999-494298/41.		
DR	P-PSDB; AAY28996.		
PT	Nuclear factor kappa B activation inhibitors, useful as preventives		
XX	for, e.g. autoimmune diseases		
XX			
PS	Examples; Page 35-39; 49pp; Japanese.		
CC	The invention provides a method for identifying or screening a nuclear		
CC	factor kappa B (NF-kB) activatin inhibitor by examining the effect of a		
CC	test substance on modulating the function(s) of TGF-beta activated kinase		
CC	1 (TAK1). The NFkB activation inhibitors targeting on TAK1 can be used to		
CC	treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis)		
CC	intractable diseases with inflammation (such as atrophic dermatitis and		
CC	psoriasis), viral infection, endotoxin shock, septicemia and others. The		
CC	nucleic acid sequence represents the nucleotide sequence of human TAK1a		
CC	(hTAK1a) protein.		
XX			
SQ	Sequence 2785 BP; 827 A; 565 C; 640 G; 753 T; 0 other;		
Alignment Scores:	2.32e-216	Length:	2785
Prod. No.:	3006.00	Matches:	578

[illegible]

QY 181 Aspl1e1glnThrHismetThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGluVal 200
 DB 697 GACATCCAAACACACATGACCAATAATAAGGAGGAGTGCTGATGGCGCTGAAAGT 756
 QY 201 PhcGluGlySerAsnTyrSerGluLysCysAspValPheSerTrpGlyLeuLeuTrp 220
 DB 757 TTTGAAGGTAGCAATTACAGTGAAGAGTGAATGCTTCAGCTGGGGATTTATCTCTGG 816
 QY 221 GluVal11eThrArgArgLysProPheAspGlu11eGlyGlyProAlaPheArg11eMet 240
 DB 817 GAAGTGATTAACAGCCGGAAACCTTCGATGAGATCGGTGGCCACCTTCGAATCATG 876
 QY 241 TrpAlaValHisAsnGlyThrArgProProLeu11eLysAsnLeuProLysPro11eGlu 260
 DB 877 TGGGCTGTTCAATATGACATCGACCACTGATCAAAATTTTACCTTAAGCCCATTTGAG 936
 QY 261 SerLeuMetThrArgCysTrpSerLysAspProSerGluAlaArgProSerMetGluGlu 280
 DB 937 AGCTTGATGACAGCTGTGTGTCTAAGGACCATCTCAGGCGCTTCAAATGGAGAAATT 996
 QY 281 ValLys11eMetThrHisLeuMetArgTyrPheProGlyValAspGluProLeuGlnTyr 300
 DB 997 GTGAAATATATGACTACTGATGCGGTACTTCCAGAGCGGATGAGCATTAACGTAT 1056
 QY 301 ProCysGlnTyrSerAspGluGlyGlnSerAsnSerAlaThrSerThrGlySerPheMet 320
 DB 1057 CTTGTGATGACTCTGTATGAAAGGAGAGCACTCAGCCACGACAGGCTGTTCATG 1116
 QY 321 Asp11e1e1eSerThrAsnThrSerAsnLysSerAspThrAsnMetGluGlnValProAla 340
 DB 1117 GACATTGCTTCTCAAAATACAGAAATAAAGTGAACAATATGGAACAGGTTCTGCGC 1176
 QY 341 ThrAsnAspThr11eLysArgLeuGlySerLysLeuLeuLysAsnGlnAlaLysGlnGln 360
 DB 1177 ACAAAACGACACTATTAAGCCTTGAGTCAAACTGTGAAAAACAGAGAAAGCAACAG 1236
 QY 361 SerGluSerGlyArgLeuSerLeuGlyAlaSerHisGlySerSerValGluSerLeuPro 380
 DB 1237 AGTGAATCTGAGCGCTGAGCTTGGGAGCCTCTGTTGGGAGCAGTGTGGAGACTTGGCC 1296
 QY 381 ProThrSerGluGlyLysArgMetSerAlaAspMetSerGlu11eGluAlaArg11eAla 400
 DB 1297 CCCACTTCCAGGAGCAAGAGATGAGTCTGACATGCTGAATAAGCCAGAGATCGTG 1356
 QY 401 AlaThrThrGlyAsnGlyGlnProArgArgArgSer11eGlnAspLeuThrValThrGly 420
 DB 1357 GCGACTGCAGTAAACGGGCAACAGCGCTGATCCATCCAAAGCTTGACTGTGCTGGG 1416
 QY 421 ThrGluProGlyGlnValSerSerArgSerSerSerProSerValArgMet11eThrThr 440
 DB 1417 ACAGAACTGTGACAGTGAACAGCCGGTCAATCCAGCCCTTATGTAATGATCACTACC 1476
 QY 441 SerGlyProThrSerGlyLysProThrArgSerHisProThrThrProAspAspSerThr 460
 DB 1477 TCGAGGACCAACCTCAGAAAGCCAGCTCGCATCCCAATGAGCCCTGTGATGATCAACA 1536
 QY 461 AspThrAsnGlySerAspAsnSer11eProMetAlaTyrLeuThrLeuAspHisGlnLeu 480
 DB 1537 GACACCAATGGCTCAGTAATCTCATCCCAATGGCGTATTTTCACTGGATCCAGCGTAA 1596
 QY 481 GlnProLeuAlaProCysProAsnSerLysGluSerMetAlaValPheGluGlnHisCys 500
 DB 1597 CAGGCTCTAGCGCGTGCCCAAACCTCCAAAGAAATCCATGCACTGTTCGAACAGCACTGT 1656
 QY 501 LysMetAlaGlnGlyUryrMetLysValGlnThrGlu11eAlaLeuLeuLeuGlnArgLys 520
 DB 1657 AAAATGGCAAGAGATATGAAAGTTCAAAACGAAATCGCATGTTACTACAGAGAAAG 1716
 QY 521 GlnGluLeuValAlaGluLeuAspGlnAspGluLysAspGlnGlnAsnThrSerArgLeu 540
 DB 1717 CAAGCACTAGTTGCAGAATTGACACGAGATGAAAAGACGACGAAATATCATCTCGTCTG 1776

QY 541 ValGlnGlnHisLysLysLeuLeuAspGluAsnLysSerLeuSerThrTyrTyrGlnGln 560
 DB 1777 GTACAGGAACATAAAGCTTTATGATGAAAACAAAGCCTTTCTATATATTCACAGCAA 1836
 QY 561 CysLysLysGlnLeuGlnVal11eArgSerGlnGlnGlnLysArgGlnGlyThrSer 579
 DB 1837 TGCAAAACCACTAGAGCTATCATCAGAACCAACAGCAAGAAACGACAGCACTTCA 1893

RESULT 8
 AAX9697
 ID AAX9697 standard; cDNA to mRNA; 2866 BP.
 XX
 AC AAX9697;
 DT 18-OCT-1999 (first entry)
 XX
 DE Human TGF-beta activated kinase (TAK) 1b encoding nucleotide sequence.
 XX
 KW Nuclear factor kappa B; NF-kB; inhibitor; TGF-beta activated kinase 1;
 KW TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;
 KW intractable disease; atrophic dermatitis; psoriasis; viral infection;
 KW endotoxin shock; septicemia; human; hTAK1b; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 163..1983
 FT FT /*tag= a
 FT FT /product= "hTAK1b"
 XX
 PN MO9940202-A1.
 XX
 PD 12-AUG-1999.
 XX
 PF 02-FEB-1999; 99MO-JP00422.
 XX
 PR 30-OCT-1998; 98JP-0309316.
 PR 06-FEB-1998; 98JP-0026003.
 XX
 PA (TANA) TANABE SEIYAKU CO.
 XX
 PI Hasegawa K, Kageyama N, Sakurai H, Sugita T;
 XX
 DR WPI; 1999-494298/41.
 DR P-PSDB; AAY28997.
 XX
 PT Nuclear factor kappa B activation inhibitors, useful as preventives
 PT for, e.g. autoimmune diseases
 XX
 PS Examples; Page 39-43; 49pp; Japanese.
 XX
 CC The invention provides a method for identifying or screening a nuclear
 CC factor kappa B (NF-kB) activation inhibitor by examining the effect of a
 CC test substance on modulating the function(s) of TGF-beta activated kinase
 CC 1 (TAK1). The NFkB activation inhibitors targeting on TAK1 can be used to
 CC treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis),
 CC intractable diseases with inflammation (such as atrophic dermatitis and
 CC psoriasis), viral infection, endotoxin shock, septicemia and others. The
 CC present sequence represents the nucleotide sequence of human TAK1b
 CC (hTAK1b) protein.
 XX
 SQ Sequence 2866 BP; 848 A; 588 C; 656 G; 774 T; 0 other;
 XX

Alignment Scores:
 Pred. No.: 1,41e-214 Length: 2866
 Score: 2982.50 Matches: 578
 Percent Similarity: 95.38% Conservative: 0
 Best Local Similarity: 95.38% Mismatches: 1
 Query Match: 98.95% Indels: 27
 DB: 20 Gaps: 1

US-09-830-144-2 (1-579) X AAX9697 (1-2866)

Qy 1 MetSerThrAlaSerAlaAlaSerSerSerSerSerAlaGlyGluMetIleGlu 20
| | | | |
Db 163 ATGCTACAGGCTCTGCCGCCCTCTCTCTCTCGTCTCGGCGGTGAGATCGAA 222
| | | | |
Qy 21 AlaProSerGlnValLeuAsnPheGluGluLeuAspTyrIysGluIleValGluGlu 40
| | | | |
Db 223 GCCCCTTCCCAGGTCCTCAACTTTGAAGAGATCGACTACAAGGAGATCGAGGTGGAAGAG 282
| | | | |
Qy 41 ValValGlyArgGlyAlaPheGlyValValCysIysAlaIysTrpArgAlaLysAspVal 60
| | | | |
Db 283 GTTGTGGAGAGAGAGGCTTTGGAGTTGTTTGCAAAGCTAAGTGGAGAGCAAAAGATGTT 342
| | | | |
Qy 61 AlaIleLysGlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
| | | | |
Db 343 GCTATTAAACAAATAGAAAGTAATCTGAGAGGAACGGTTTATTGTAGAGCTTTCCGCAG 402
| | | | |
Qy 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnProVal 100
| | | | |
Db 403 TTATCCCGTGTGAACCATCTTAATATTGTAAGCTTTATGGAGCCTGCTTGAATCCAGTG 462
| | | | |
Qy 101 CysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGlyAlaGlu 120
| | | | |
Db 463 TGTCTTGATGGAAATATGCTGAAGGGGGCTCTTTATATAATGTGCTGCATGGTGTGAA 522
| | | | |
Qy 121 ProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeuGlnCysSerGlnGly 140
| | | | |
Db 523 CCATTGCCATATTATATCTGTGCCCGCAATGAGTTGGTGTTCACAGTGTCCCAAGGA 582
| | | | |
Qy 141 ValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysProPro 160
| | | | |
Db 583 GTGGCTTATCTTACAGCATGCAACCCCAAGCGCTAATTCACAGGACCTGAAACCA 642
| | | | |
Qy 161 AsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCys 180
| | | | |
Db 643 AACTTACTGTGTGTGCAGGGGGACAGTCTTAAAAATTTGTGATTTTGTGTACAGCCTGT 702
| | | | |
Qy 181 AspileGlnThrHisMetThrAsnAsnLysGlySerAlaIleTrpMetAlaProGluVal 200
| | | | |
Db 703 GACATTGAGACACACATGACCAATAACCAAGGGGAGTCTGCTTGGATGGCACCTGAAGTT 762
| | | | |
Qy 201 PheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrpGlyIleIleLeuTrp 220
| | | | |
Db 763 TTTGAAGGTAGTAAATACAGTGAANAATGTGACGTCTTCAGCTGGGGTATTATCTTTGG 822
| | | | |
Qy 221 GluValIleThrArgArgLysProPheAspGluIleGlyProAlaPheArgIleMet 240
| | | | |
Db 823 GAAGTGATAACGCTCGAAACCCCTTTGATGAGATTGTTGCCCCAGCTTTCGAATCATG 882
| | | | |
Qy 241 TrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGlu 260
| | | | |
Db 883 TGGGCTGTTTATAATGTGTACTCGACCACTCTGATAAAAAATTTACCTAAGCCCATTTGAG 942
| | | | |
Qy 261 SerLeuMetThrArgCysTrpSerLysAspProSerGlnArgProSerMetGluGluIle 280
| | | | |
Db 943 AGCCTGATGACTCGTTGTTGTTAAAGATCCCTTCCAGCGCCCTTCAATGGAGGAAAT 1002
| | | | |
Qy 281 ValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyr 300
| | | | |
Db 1003 GTGAAAATAATGACTCACTTGATGCGGTACTTCCAGGACAGATGAGCCATTACAGTAT 1062
| | | | |
Qy 301 ProCysGlnTyrSerAspGluGlyGlnSerAsnSerAlaThrSerThrGlySerPheMet 320
| | | | |
Db 1063 CCTTGTGAGTATTGAGATGAAGACAGACGAACTCTGCCACCACTGACAGGCTCATTCATG 1122
| | | | |
Qy 321 AspileAlaSerThrAsnThrSerAsnLysSerAspThrAsnMetGluGlnValProAla 340
| | | | |
Db 1123 GACATTGCTTCAAAATACGAGTAACAAAGTGACACTAATATGAGCAAGTTCCTGTC 1182
| | | | |
Qy 341 ThrAsnAspThrIleLysArgLeuGluSerLysLeuLysAsnGlnAlaLysGlnGln 360
| | | | |
Db 1183 ACAATGATACTATTAGCGCTTAGAATCAAAATGTTGAAAATAATCAGGCAAGCAACAG 1242
| | | | |
Qy 361 SerGluSerGlyArgLeuSerLeuGlyAlaSerHisGlySerSerValGluSerLeuPro 380
| | | | |

Db 1243 AGTGAATCTGGACGTTTAAAGCTTTGGAGCCTCCCTGGGAGCAGTGTGGAGAGCTTGCCC 1302
| | | | |
Qy 381 ProThrSerGluGlyLysArgMetSerAlaAspMetSerGluIleGluAlaArgIleAla 400
| | | | |
Db 1303 CCAACCTCTGAGGGCAAGAGGATGAGTCTGNACTGTCTGAAATAGAAAGCTAGGATCGCC 1362
| | | | |
Qy 401 AlaThrThr----- 403
| | | | |
Db 1363 GCAACACAGCCTATTCCAAGCCTAAACGGGCCACCCTGAAAACTGCTTCTTATTTGGCAAC 1422
| | | | |
Qy 404 -----GlyAsnGlyGlnProArgArgArgSerIle 413
| | | | |
Db 1423 ATTCTGGAGTGTCCCTGAGATCGTCATATCAGCAACGACAGCCCAAGAGCTAGATCCATC 1482
| | | | |
Qy 414 GlnAspLeuThrValThrGlyThrGluProGlyGlnValSerSerArgSerSerSerPro 433
| | | | |
Db 1483 CAAGACTTGACTGTAACTGGAAACAGAACTGGTCAGGTGAGCAGTAGGTCTATCCAGTCCC 1542
| | | | |
Qy 434 SerValArgMetIleThrThrSerGlyProThrSerGluLysProThrArgSerHisPro 453
| | | | |
Db 1543 AGTGTGAGAAATGATTACTACCTCAGGACCAACCTCAGAAAAAGCCAACTCGAAGTCATCCA 1602
| | | | |
Qy 454 TrpThrProAspAspSerThrAspThrAsnGlySerAspAsnSerIleProMetAlaTyr 473
| | | | |
Db 1603 TGGACCCCTGATGATTCACAGATACCAATGGATCAGATAACTCCATCCCAATGGCTTAT 1662
| | | | |
Qy 474 LeuThrLeuAspHisGlnLeuProLeuAlaProCysProAsnSerLysGluSerMet 493
| | | | |
Db 1663 CTTACACTGGATCACCACCTACAGCCTCTAGACCGTCCCAAACTCCAAAGAATCTATG 1722
| | | | |
Qy 494 AlaValPheGluGlnHisCysLysMetAlaGlnGluTyrMetLysValGlnThrGluIle 513
| | | | |
Db 1723 GCAGTGTTTGAACAGCAATTGTAAAAATGSCACAGAATATATGAAAGTTCAACAGAAAT 1782
| | | | |
Qy 514 AlaLeuLeuLeuGlnArgLysGlnGluLeuValAlaGluLeuAspGlnAspGluLysAsp 533
| | | | |
Db 1783 GCATTGTTATTACAGAGAAACAGAACTAGTTGCAGAACTGGACCAGGATGAAAGGAC 1842
| | | | |
Qy 534 GlnGlnAsnThrSerArgLeuValGlnGluHisLysLysLeuLeuAspGluAsnLysSer 553
| | | | |
Db 1843 CAGCAAAATACATCTCGCCTGTGACAGGAACATAAAAAAGCTTTTAGATGAAAAACAAAGC 1902
| | | | |
Qy 554 LeuSerThrTyrTyrGlnGlnCysLysLysGlnLeuGluValIleArgSerGlnGln 573
| | | | |
Db 1903 CTTTCTACTTACTACCACCAATGCAAAAAACAACCTAGAGGTCTATCAGAGTCAGCAGCAG 1962
| | | | |
Qy 574 LysArgGlnGlyThrSer 579
| | | | |
Db 1963 AAACGACAAAGGCACCTTCA 1980
| | | | |
RESULT 9
AAx99698
ID AAx99698 standard; cDNA to mRNA; 1704 BP.
XX
AC AAx99698;
XX
DT 18-OCT-1999 (first entry)
XX
DE Human TGF-beta activated kinase (TAK) 1c encoding nucleotide sequence.
XX
KW Nuclear factor kappa B; NF-kB; inhibitor; TGF-beta activated kinase 1;
KW TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;
KW intractable disease; atrophic dermatitis; psoriasis; viral infection;
XX endotoxin shock; septicemia; human; hTAK1c; ss.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1704
FT /*tag= a
FT /product= "hTAK1c"
XX

QY 494 AlaValPheGluGlnHisCysLysMetAlaGlnGluTyrMetLysValGlnThrGluIle 513
Db 1524 ----- 1524
QY 514 AlaLeuLeuGlnArgLysGlnGluValAlaGluLeuAspGlnAspGluLysAsp 533
Db 1525 -----CAAGAACTAGTTGCGAAGCTGACAGGATGAAAGGAC 1563
QY 534 GlnGlnAsnThrSerArgLeuValGlnGluHisLysLysLeuAspGluAsnLysSer 553
Db 1564 CAGCAAAATACATCTCGCCTGGTACAGGACATAAAGCTTTTATGATGAAACAAAGGC 1623
QY 554 LeuSerThrTyrThrGlnGlnCysLysLysGlnLeuGluValIleArgSerGlnGlnGln 573
Db 1624 CTTTCTACTTACTACAGCAATGCAAAACAACTAGAGGTCTATCAGAGTCTCAGCAGCAG 1683
QY 574 LysArgGlnGlnThrSer 579
Db 1684 AAACGAAAGGCACCTTCA 1701

RESULT 10

ABL02489

ID ABL02489 standard; cDNA; 3367 BP.

XX ABL02489;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 1949.

DE Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical; gene; ss.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX P-FSDB; ABB58386.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

XX genes from Drosophila and for elucidating cell signalling and cell-cell

XX interactions -

XX Claim 1; SEQ ID NO 1949; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

XX capable of detecting 1000 or more genes from Drosophila. The invention is

XX useful in developmental biology and in elucidating cell signalling and

XX cell-cell interactions in higher eukaryotes for the development of

XX insecticides, therapeutics and pharmaceutical drugs. The invention

XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

XX sequences (ABL01840-ABL16175) and the encoded proteins

XX (ABB57737-ABB72072).

XX The sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 3367 BP; 938 A; 898 C; 856 G; 675 T; 0 other;

XX Alignment Scores:

Pred. No.: 2,66e-63 Length: 3367
Score: 971.00 Matches: 232
Percent Similarity: 52.22% Conservative: 97
Best Local Similarity: 36.83% Mismatches: 202
Query Match: 32.22% Indels: 100
DB: 23 Gaps: 14

US-09-830-144-2 (1-579) x ABL02489 (1-3367)

QY 17 GluMetIleGluAlaProSerGlnValLeuAsnPheGluGluIleAspTyrLysGluIle 36
Db 932 GAGATGCCACAGCATCGCTGGACGCACATCGAGCAGCCTATGTGGACTTTCAGTGAGATA 991
QY 37 GluValGluGluValValGlyArgGlyAlaPheGlyValValCysLysAlaLysTyrArg 56
Db 992 ACATAAGAGAGAAAGTCGGCCATGGGTCTCAGGAGTGTCTGCAAGGCCGTTTGGCGC 1051
QY 57 AlaLysAspValAlaIleLysGlnIleGluSerGluArgLysAlaPheIleVal 76
Db 1052 GACAAGCTGGTTGCCGTCAAGGAGTTCTTCGCCAGCGCCGAGCAGGACATCGAGAAG 1111
QY 77 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGly----- 94
Db 1112 GAGGTGAAGCAGTGTCTCGCGCTGAAGCACCCGAAACATCATCTCTGCACGGGATATCC 1171
QY 95 AlaCysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsn 114
Db 1172 TCGTACCAGCAGGCCACCTACCTAGTAATAGAGTTCGCCAAGGTGGATCGCTGCAAC 1231
QY 115 ValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCys 134
Db 1232 TTCCTTCACGCG--AAGGTGAAGCGCGCATATTCCTGGCCCCACGCCATGAGCTGGCG 1288
QY 135 LeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHis 154
Db 1289 CGCCATGTGCAGAGGGTCTGGCATATTTGCATGCCATGACGCCAAACCACTAATACAT 1348
QY 155 ArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCys 174
Db 1349 CGCGACGTGAAGCGCTGAACCTGCTTTGACCAACAGGAGCGCAATCTGAAGATATGC 1408
QY 175 AspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAla 194
Db 1409 GACTTCGGCACGGTGGCGACAAGTCGACCATGATGACCAACAATCGCGCAGTGGCGT 1468
QY 195 TrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSer 214
Db 1469 TGGATGGCGCCCGAGGTCTCGAAGGCTCCAAGTATACGAGAAGTGTGACATTTTATAG 1528
QY 215 TrpGlyIleIleLeuTrpGluValIleThrArgAlaGlyLysProPheAspGluIleGly 234
Db 1529 TGGGCCATTGTTCTATGGGAGGTTCGTCCAGGAAGCAGCCCTTTAAAGGCATCGCAAT 1588
QY 235 ProAlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsn 254
Db 1589 ---GCCTACACCATCCAGTGAAGATCTACAAGGGTGAACGCCCGCGCTGTGACCACT 1645
QY 255 LeuProLysProIleGluSerLeuMetThrArgCysTyrSerLysAspProSerGlnArg 274
Db 1646 TGCCCCAAGCGCATCGAGACCTGATGACCGCTGCTGGAAACACGTCGCCGAGGATCGC 1705
QY 275 ProSerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAla 294
Db 1706 CCGTCGATCAGTACATAGTGGCGTTATGCAGGATCGTCAAGACTATATACGGGGCGG 1765
QY 295 AspGluProLeuGlnTyr----- 300
Db 1766 GACAAGGCCCTGGAATACACGTTTGTATCAACAGATTGTCAACAAAGAGACGACGCGC 1825
QY 301 -----ProCysGlnTyrSerAspGluGlyGlnSerAsn----- 311
Db 1826 ACGGTGGCGCGCTCAACCGGATAGCCTCAGTTCGAGGAGGGGGAACATGAGCCCTCGTCC 1885

OY 312 -----SerAla 313
Db 1886 ACACAGTTAACCGACAACGGCGCCAAACGCCAATGTGAACGCGATGACATATCAAAA 1945
OY 314 ThrSerThrGlySerPheMetAspIleAlaSerThrAsnThrSerAsnLysSerAspThr 333
Db 1946 ACACAGACTAGCTCAATGACCGAATAATCTTCATCATCATCGACATCGACATCGACGCG 2005
OY 334 AsnMetGluGlnValPro-----AlaThrAsn-----AspThr 344
Db 2006 AACTCGGGCCACTGGACATATATCCGCTATTCTTATATGTCACCAATCGCTGGACCGC 2065
OY 345 IleLysArgLeuGluSerLysLeuLysAsn----- 355
Db 2066 ATTTCCCGAGAGAGAGAGACGAGCGGACGAGTAGCTTCACCTGACCTGCTGCGCT 2125
OY 356 GlnAlaLysGlnGlnSerGluSerGlyArgLeuSerLeuGlyAlaSerHisGlySerSer 375
Db 2126 GAGGCCACTACGCGCTCGAAAAGCATCGGAAAGCATGATCTGATGCGCTGCAAGCCC 2185
OY 376 ValGluSerLeuProProThrSerGluGlyLysArgMetSerAlaAspMetSerGluIle 395
Db 2186 ATGGAGCACTC-----ACCTTCACGTCGAGCGGAATGCGTTGATGTGATGCCAGC 2239
OY 396 GluAlaArgIleAlaAlaThrThrGlyAsnGlyLysProArgArgSerIleGlnAsp 415
Db 2240 GAAAGCAGCAGCAGCAGCAGCAAGCAAGCAGATGCGCGCAAGCA----- 2287
OY 416 LeuThrValThrGlyThrGluPro----- 423
Db 2288 CTCACCGTACGACGACCAAGCGGTGATGATGATGACCCAGGATCTGCCAACAAACAGGC 2347
OY 424 Gly-GlnValSerSerArgSerSerSerProSerValArgMetIleThrThrSerGlyPr 443
Db 2348 GGCATCCAGCCGACCTCGAAGCAGACGCGCATGCGCATGCGAATGGTT-----GGCA 2398
OY 443 oThrSerGluLysProThrArgSerHisPro-----TrpThr 455
Db 2399 GCMAAGATGAGAGAGCTGACGAGGCAAGAGCATGACAGAGATTTGTCAACTCGTTGGAC 2458
OY 455 rProAspArgSerThrAspThrAsnGlySerAspAsnSerIleProMetAlaTyrLeuTh 475
Db 2459 GTC-GACGTGGATCCGACAGATGAGACGAGCGACCGAAGACTACTGCGCGAGAT 2517
OY 475 rLeuAspHisGlnLeuGlnProLeuAlaProCysProAsnSerLysGluSerMetAlaVal 495
Db 2518 TCTTGATCCGAGACTCCAGCAGCGCGGATATCCCAAGATGCGGATCCAGCTCAT 2577
OY 495 lPheGluGlnHisCysLysMetAlaGlnGluTyrMetLysValGlnThrGluIleAlaLe 515
Db 2578 CTACCGGAGACCGACGACATGCGCAAGAGTCTGACGCTGACACAGCAACCTCTACTA 2637
OY 515 uLeuLeuGlnArgLysGlnGlnLeuValAlaGluLeuAspGlnAspGluLysAspGlnI 535
Db 2638 CGCGCAGAGACTTAAAGACAGACTCATCGTCAGATGAGACCGAAGCGGAAACAGAA 2697
OY 535 nAsnThrSerArgLeuValGlnGlnHisLysLysLeuLeuAspGluAsnLysSerLeuSe 555
Db 2698 G-----CAGGAGCTTTCGCGCAAGATGAAGACAAAGAGGGTCTTCA 2739
OY 555 rThrTyrTyrGlnGlnCysLysLysGln 564
Db 2740 GAGTCTTTACACATCTGCAGCAGAG 2767
RESULT 11
ID ABL02488 standard, cDNA, 10997 BP.
XX ABL02488;
XX 26-MAR-2002 (first entry)
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 1946.

XX KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX OS Drosophila melanogaster.
XX PN MO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US09231.
XX PR 23-MAR-2000; 2000US-191637P.
XX PR 11-JUL-2000; 2000US-0614150.
XX PA (PEKE) PE CORP NY.
XX PI Venter JC, Adams M, Li PMD, Myers EW;
XX DR WPI, 2001-656860/75.
XX DR P-PSDB; ABB58385.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX PT interactions -
XX PS Claim 1; SEQ ID NO 1946; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
XX CC sequences (AB16176-AB16175) and the encoded proteins
XX CC (ABB5737-ABB72072).
XX CC The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX CC
SQ Sequence 10997 BP; 3237 A; 2362 C; 2382 G; 3016 T; 0 other;
Alignment Scores:
Pred. No.: 1,84e-42 Length: 10997
Score: 702.50 Matches: 190
Percent Similarity: 41.09% Conservative: 73
Best Local Similarity: 29.69% Mismatches: 164
Query Match: 23.31% Indels: 214
DB: 23 Gaps: 14
US-09-830-144-2 (1-579) x ABL02488 (1-10997)
OY 25 ValLeuAsnPhelGluGlnIleAspTyrLysGluIleGluVal----- 38
Db 2469 GTAATTGATATTGAGAGATTTCGACAAAGAGGTTACACTGTTGATCAGATCTATAA 2528
OY 39 -----GluGluValValGlyArgGlyAlaPheGlyValValCysLysAlaLysTrp 55
Db 2529 TCTTTTCCCGCAAAAGTCGCGCATGGCTCTTACGGAATGGCTGCAAGGCGCTTGG 2588
OY 56 ArgAlaLysAspValAlaIleLysGlnIleGluSerLysSerGluArgLysAlaPheIle 75
Db 2589 CGCGACAGAGCTGGTTCGCTTAAGAGTTCTTCCGACGCCCGGACGAAAGACATCGAG 2648
OY 76 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGly--- 94
Db 2649 AAGAGGTGAAGAGAGTTGTCGCGGTGAAGACCCGAAACATCATCGCTGACGCGGATA 2708
OY 95 ---AlaCysLeuAsnProValCysLeuValMetGluTyrTrpAlaGlnGlyLysSerLeuTyr 113
Db 2709 TCCTGTACACAGAGGCCACTTACTGATTAATGAGTTCCGCGAAGGTGATCGCTGCAC 2768
OY 114 AsnValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrp 133

CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABBS7737-ABBS72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 759 BP; 226 A; 162 C; 198 G; 173 T; 0 other;

Alignment Scores:

Pred. No.:	4,16e-29	Length:	759
Score:	505.00	Matches:	99
Percent Similarity:	60.25%	Conservative:	48
Best Local Similarity:	40.57%	Mismatches:	71
Query Match:	16.76%	Indels:	26
DB:	23	Gaps:	4

US-09-830-144-2 (1-579) x ABL08337 (1-759)

```
QY 29 GUGLUILASPTLYRSGIULIEGLUVALIGLUVALIGLYARGGLYALAPHEGLY 48
    |||  ::  :::::|||||:  :::::|||||:  :::::|||||:  :::::|||||:
Db 16 GAGGAGTTCCTTATGAGGATCCAGACAAAGAGCTTATGAGCCGAGTTCTATGGC 75
QY 49 VALVALCYELYSALALYSTTPARGALALYSASPVALALILELGSGLINLEGLUSERGLU 68
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 76 AGCGTATACAGGCGGCTGCGAATAATCGTAGATCGCCCTGAGAGGATTCGACAGGA 135
QY 69 SERGLARGLYSALAPHELLEVALIGULEUARGINLEUSERARGVALASNHISPROASN 88
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 136 TGTGAGAGATTAATAAGTTGAAAGAGAGATCTACACACTTACAAAGCCCATGTCAAC 195
QY 89 ILEVALYSLEUTYRIGLYALACYSLEUASPROVALYS-----LEUVALMETGLUTYR 106
    |||  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 196 ATAGTTGAGACTCTACGCGCATCGACAGGAGGATGCGCCCTGTTGATGGAATTC 255
QY 107 ALIAGLUGLIGLYSERLEUTYRASHVALLEUHSGLYALIGLUPROLEUPROTYRTHYR 126
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 256 GTAGAGGCGGATCTCTGTCAGATTTCGCAC---GCCAAAAGCAAGCAGATTATTCG 312
QY 127 ALAALAHIALAMETSERTPCYSLLEUGINCYSERINGLIVALAITYRLEUHSER 146
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 313 CATGCCCAAGCTTCAACACTGGCGCATCATGCTCTCAGGAGATGCTATTCGATGCC 372
QY 147 METGLINPROLYSALALEUILEHISARGASPLEULYSPROPROASNLEULEUVALA 166
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 372 ATGCAAGCCGAAAGCAGATATTCATCGCATTAAGCCACTCAATACACTGCTATGCGAG 432
QY 167 GLYGLYTHYRVALLEULYSILECYASAPHEGLYTHYRALIACYSASPILEGLINTHRISMET 186
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 433 AAGGAGCTCAAACTGAAGATTGCGATTTCGAACTGTTGTGAGACTTATCCCAATGCATA 492
QY 187 THRASNASHLYSGLYSERIALALATRPMETALAPROGLU----- 199
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 493 TCGTGCAATGCGGCGACCTCGCATACAAAGGCCGCCGAGTAAGGAGACTGTTGATTC 552
QY 200 -----ValpHEGLUGLYSER 204
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 553 AATATCAATCGAATAATATATCAACCAACCAACCGGCTTCAAAAGGTTCTACAGGAAT 612
QY 205 AANTYRSEGLULYSYASAPVALAPHESETTPGLYILELEUTRGLUVALILEHR 224
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 613 AAACCCGATGAAGAAGTGGAGTGTATGTTGGCTATTAACCTTTGGAAATATGTGCG 672
QY 225 ARGARGLYSPROPHASPLULEGLYGLYPROALAPHEARGILEMETTRPALVALHS 244
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 673 CGCAGAGAGCCATTGACAGATAT---AATACGCTTTTGAACGTACATGCGCTATTTAT 729
QY 245 ASNGLYTHYRARG 248
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 730 GAAGGCAAGAGA 741
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
```

RESULT 13

ABL70018

ID ABL70018 standard; DNA; 3454 BP.

XX ABL70018;

AC 15-MAY-2002 (first entry)

DE Pancreas cancer related gene sequence SEQ ID NO:8355.

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
XX stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
XX cytosolic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
XX gene; ds.

OS Homo sapiens.

XX WO200194629-A2.

XX 13-DEC-2001.

PF 30-MAY-2001; 2001WO-US10838.

XX 05-JUN-2000; 2000US-209473P.

PR 05-JUN-2000; 2000US-209531P.

PR 18-SEP-2000; 2000US-233133P.

PR 18-SEP-2000; 2000US-233617P.

PR 20-SEP-2000; 2000US-234009P.

PR 20-SEP-2000; 2000US-234034P.

PR 20-SEP-2000; 2000US-234052P.

PR 22-SEP-2000; 2000US-234509P.

PR 22-SEP-2000; 2000US-234567P.

PR 25-SEP-2000; 2000US-234923P.

PR 25-SEP-2000; 2000US-234924P.

PR 25-SEP-2000; 2000US-235077P.

PR 25-SEP-2000; 2000US-235082P.

PR 25-SEP-2000; 2000US-235134P.

PR 25-SEP-2000; 2000US-235280P.

PR 26-SEP-2000; 2000US-235637P.

PR 26-SEP-2000; 2000US-235638P.

PR 27-SEP-2000; 2000US-235711P.

PR 27-SEP-2000; 2000US-235720P.

PR 27-SEP-2000; 2000US-235840P.

PR 27-SEP-2000; 2000US-235863P.

PR 28-SEP-2000; 2000US-236028P.

PR 28-SEP-2000; 2000US-236032P.

PR 28-SEP-2000; 2000US-236033P.

PR 28-SEP-2000; 2000US-236034P.

PR 28-SEP-2000; 2000US-236109P.

PR 28-SEP-2000; 2000US-236111P.

PR 29-SEP-2000; 2000US-236842P.

PR 29-SEP-2000; 2000US-236891P.

PR 02-OCT-2000; 2000US-237172P.

PR 02-OCT-2000; 2000US-237173P.

PR 02-OCT-2000; 2000US-237278P.

PR 02-OCT-2000; 2000US-237294P.

PR 02-OCT-2000; 2000US-237295P.

PR 02-OCT-2000; 2000US-237316P.

PR 03-OCT-2000; 2000US-237425P.

PR 03-OCT-2000; 2000US-237598P.

PR 03-OCT-2000; 2000US-237604P.

PR 03-OCT-2000; 2000US-237606P.

PR 03-OCT-2000; 2000US-237608P.

PR 01-NOV-2000; 2000US-244867P.

PR 01-NOV-2000; 2000US-245084P.

PA (AVAL-) AVALON PHARM.

PI Young PE, Auguster M, Carter KC, Ebner R, Endress G, Horrigan S;

XX Sopet DR, Weaver Z;

XX WPI; 2002-188264/24.

XX Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set -
 XX
 XX Claim 1; SEQ ID 8355; 44pp; English.

XX The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytosstatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilm's tumour.

XX
 SQ Sequence 3454 BP; 594 A; 1217 C; 1136 G; 507 T; 0 other;

Alignment Scores:
 Pred. No.: 2,13e-27 Length: 3454
 Score: 493.50 Matches: 145
 Percent Similarity: 46.49% Conservative: 87
 Best Local Similarity: 29.06% Mismatches: 191
 Query Match: 16.37% Indels: 79
 DB: 24 Gaps: 14

US-09-830-144-2 (1-579) x ABL70018 (1-3454)

QY 21 AlaProSerGlnValLeuAsnPheGluGluIleAspTyrLysGluIleGluValGluGlu 40
 DB 535 GCACCCGGGGCTCCAGCTGCCCCAGGAGATCCCTTCCAGAGCTGCAGCTAGGAGG 594
 QY 41 ValValGlyArgGlyAlaPheGlyValValCysLysAlaLysTyrArgAlaLysAspVal 60
 DB 595 ATCATCGGTGTGGGGGCTTTGGCAAGGTCTATCGGGCCCTGTGGCGTGGCGAGGAGTG 654
 QY 61 AlaIleLysGlnIleGluSerGluArgLysAlaPheIleVal----- 76
 DB 655 GCAGTCAAGGCGCGCGCTGGACCTGTGAGAGGACCCGGCAGTGACAGCGGAGGAGTG 714
 QY 77 -----GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGly 94
 DB 715 TGCCAGGAGCCCGGCTCTTTGGAGCCCTGCAGACCCCAACATAATTGCCCTTAGGGGC 774
 QY 95 AlaCysLeuAsnPro-----ValCysLeuValMetGluTyrAlaGluGlySerLeu 112
 DB 775 GCCTGCTCAACCCCCACACCTCTTCCTAGTAGTATGCGGGGCTGTGCACCTG 834
 QY 113 TyrAsnValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSer 132
 DB 835 AGCAGGGTGTGGCAGGTGCGCGGGTGCCACCTCACGTG-----CTGGTCAAC 882
 QY 133 TrpCysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeu 152
 DB 883 TGGGTGTGAGGTGGCCGGGCGATGAACCTACTACACATGATGCCCTGTGCCCATC 942
 QY 153 IleHisArgAspLeuLysProProAsnLeuLeuVal----- 165
 DB 943 ATCCACGGGACCTCAAGTCCATCAACATCTCTGATCCCTGGAGGCCATCAGAACACACAC 1002
 QY 166 AlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCysAsp---IleGlnThr 184
 DB 1003 CTGCGACACACGGTGTCTCAAGATCAGGACTTCGGCTTCGGCGCGAGTGGCACACAGACC 1062

QY 185 HisMetThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGluValPheGluGlySer 204
 DB 1063 ACAACATGATGCGCTCGGGGACCTACGCTGGATGGCGCGAGGTTATCGGTCTCTCC 1122
 QY 205 AsnTyrSerGluLysCysAspValPheSerTrpGlyIleIleLeuTrpGluValIleThr 224
 DB 1123 CTCCTTCTCCAAAAGCAGTGTCTGGAGCTTCGGGGTCTGCTGTGGAGCTGTGAGC 1182
 QY 225 ArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMetTrpAlaVal--- 243
 DB 1183 GGGAGGTCCCTTACCGTGAGATC-----GACGCTTGGCGTGGCGTATGGCTGGCT 1236
 QY 244 HisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGluSerLeuMet 263
 DB 1237 ATGAATAAGCTGACGTGCGCATTCCTCCAGCTGCCCGAGCCCTTTGCCCGCCTCTG 1296
 QY 264 ThrArgCysTrpSerLysAspProSerGlnArgProSerMetGluGluIleValIle 283
 DB 1297 GAGGAATGCTGGGACCCAGACCCACGGCGGCAGATTTCGGTAGCATCTTGAAGCGG 1356
 QY 284 MetThrHisLeuMetArgTyrPhePro----- 292
 DB 1357 CTTGAAGTCAATCAACAGATC-AGCCCTGTTCCAGATGCCATGGAGTCCCTCCACTCGCT 1415
 QY 292 ----- 292
 DB 1416 GCAGGAAGCTGGAAGCTGGAGATTGAGCATGTTTGTGATGACCTTCGGACCAAGAGAA 1475
 QY 293 GlyAlaAspGluProLeuGlnTyrProCysGlnTyrSerAspGluGlyGlnSerAsnSer 312
 DB 1476 GGAGCTTCGGAGCGGTG-AGGAGGAGCTGCTGGCGGCGCAGAGGAGCAGCGCTCCAGG 1534
 QY 313 AlaThrSerThrGly-----SerPheMetAspIleAlaSerThrAsnThrSer 328
 DB 1535 AGGAGCAGCTGCGCGCGGCGAGCAGGAGTGGCAGACAGCTGAGATGGACATCTGTGAAC 1594
 QY 329 AsnLysSerAspThrAsnMetGluGlnValProAlaThrAsnAspThrIleLysAspLeu 348
 DB 1595 GGGAGCTGCACCTGCTC-ATGTCCAGCTGACGACGAGAGAGCCCGGGTCCGCAAGCGC 1653
 QY 349 GluSerLysLeuLysAsnGlnAlaLysGlnGlnSerGluSerGly-----Arg 365
 DB 1654 AAGGGCAACTTCAAGCGCAGCCCGCTGCTCAAGCTGGGGAAGCGCGCAGCCACATCAGC 1713
 QY 366 LeuSerLeuGlyAlaSerHisGlySerSerValGluSerLeuProProThrSerGluGly 385
 DB 1714 CTGCCCTCTGGCTTTGAGCATAAAGATCAGACTCCAGGCTCTCCAACTCTCTGGAT----- 1767
 QY 386 LysArgMetSerAlaAspMetSerGluIleGluAlaArgIleAlaAlaThrThrGlyAsn 405
 DB 1768 AAGCGGAAGGATCCGATGGGCGCCAGCCCTTCGCAAGCCCGCAGCATCATC----- 1818
 QY 406 GlyGlnProArgArgArgSerIleGlnAspLeuThrValThrGlyThrGluProGlyGln 425
 DB 1819 -----CCCGCGCTGAGGCGCATTCGC-----CTGACTCCCGTGGACTGTGGTGC 1863
 QY 426 ValSerSerArgSerSerProSerValArgMetIleThrThrSerGlyProThrSer 445
 DB 1864 AGCAGCAGTGCAGCAGCAGTGGAGGAAGTGGACATGGAGCGCGGTGGCGCCCAAG 1923
 QY 446 GluLys-----ProThrArgSerHisProThrThrProAspAspSer 459
 DB 1924 AAGGAAGAACTGTGCGGGGCAAGAGAGGAGCAACGTTGGGGGCCAGCTCCACC 1980
 RESULT 14
 AAZ99726
 ID AAZ99726 standard; cDNA; 2120 BP.
 XX
 XX AAZ99726;
 AC
 XX
 XX 12-JUL-2000 (first entry)
 DT
 XX cDNA encoding human cardiovascular system associated protein kinase-2.
 DE

QY	64	-----GlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln	80
Db	182	AAGCTCTCCAAATAATGAGAAAGAGCAGAA-----ATA	214
QY	81	LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnPro-----	99
Db	215	CTCAGTGTCTCAGTCCACAGAAACATATCATCCAGTTTATATGAGTAATCTTGAACCTCCC	274
QY	100	---ValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGly	118
Db	275	AACTATGGCATTTGTCACAGAAATATGCTTCTCTGGAGATCACTATGATTACATTACAGT	334
QY	119	-----AlaGluProLeuProTyrTyrThrAlaAlaHisIleMetSerTyrCysLeuGln	136
Db	335	AACAGAAAGTAGAGATG-----GATATGGATACATTATGACCTGGCCACTGAT	385
QY	137	CysSerGlnGlyValAlaTyrLeuHisSerMetGlnProValAlaLeuIleHisArgAsp	156
Db	386	GTAGCCAAAGGAATGCAATTATTTATCATATGAGAGCTCTGTCAAGTGATTTCACAGAGC	445
QY	157	LeuLysProProAsnLeuLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPhe	176
Db	446	CTCAAGTCAGAAAAGCTGGTTATATAGCTGCATAGA---GTACTGAAGATCTGTGACTTT	502
QY	177	GlyThrAlaCysAspIleGlnThrHisMetThrAsnLys-----GlySerAlaAla	194
Db	503	GGT---GCCCTGGTTCCTCAATACCAACACACATATGCTTGGTGAACCTTTCCCA	559
QY	195	TrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSer	214
Db	560	TGGAGGCTCCAGAAAGTTATCCAGAGCTCCCTGTGTCAAAAATTGTGACACATATTCC	619
QY	215	TyrGlyIleIleLeuTyrProGluValIleThrArgArgLysProPheAspIuIleGlyGly	234
Db	620	TATGCTGTGCTCTCTGTGGAGATCTTACAGAGAGCTCCCTTTAAAGCTTTGAGAGA	679
QY	235	ProAlaPheArgIleMetTyr--AlaValHisAsnGlyThrArgProProLeuIleLys	253
Db	680	-----TTACAAAGTAGCTGGCTTGTATAGTGGAAAAAAGAGAGATTACCATTCACAGC	733
QY	254	AsnLeuProLysProIleGluSerLeuMetThrArgCysTyrSerIleAspProSerGln	273
Db	734	AGTTGCCCCAGAAAGTTTGTCTGACATGTTTACATCACTGTGTGGAAAGCTGATCCCAAGAA	793
QY	274	ArgProSerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGly	293
Db	794	CGGCGATCATTTCAAGCAAAATCATTTCAATCCGT-----	826
QY	294	AlaAspIuProLeuGlnTyrProCysGlnTyrSerAspGluGlyGlnSerAsnSerAla	313
Db	827	-----GACTCCATGTCAAAATACAGC	847
QY	314	ThrSerThrGlySerPheMetAspIleAlaSerThrAsnThrSerAsnLysSerAspThr	333
Db	848	-----AGCCTTCTGTGACAGTGAATCTCATCTTCTACAAACAAAGCCGAGTGC	895
QY	334	AsnMetGluGlnValProAlaThrAsnAspThrIleLysArgLeuGluSerLysLeuLeu	353
Db	896	AGGTGCCAA---ATTAGGCAACTCTTGAGAGGCTTAAAGAACTTGACGCGATCTCAGC	952
QY	354	LysAsnGlnAlaLysGlnGlnSerGluSerArgLysGluSerLeu-----GlyAla	370
Db	953	TTTAAAGAGAGCAGAGCTTAAAGAACAGAAACAGTTTAAAGATGTGGACCAAAACCTG	1012
QY	371	SerHisGlySerSerValGluSerLeuProProThrSer-----	383
Db	1013	ACAGAGCAGTCCAAACCCCGCTTCTCTTGCTTGCTGCAAGAAATGTGAGAGAGCT	1072
QY	384	-----GluGlyLysArgMetSerAlaAspMetSerGluIleGluAlaArgIleAla	401
Db	1073	TACTTTGAATCTTAAACAGAGAGATCAAAACAGTGCAGAGATGTCAATGATCAACGCA	1132
QY	402	ThrThr-----GlyAsnGlyIuProArgArgArgSerIleGlnAspLeuThrValThrGly	420

```
||||: ||| ||| |||: ||| |||: ||| |||: |||
Db 1133 ACAAGTAACGGGAGGCCATGGCATGAACCAAGTCTGCAGGCCATGATGCTGATGGC 1192
Qy 421 ThrGluProGlyGlnValSerSerArgSerSerProSerVal 435
Db 1193 TTT-----GGGATATCTTCTCAATGAACAAAGCAGGAGCTGTG 1231
RESULT 15
AAA75674
ID AAA75674 standard; DNA; 1706 BP.
XX
AC AAA75674;
XX
DT 22-JAN-2001 (first entry)
XX
DE DNA encoding a human regulator of intracellular phosphorylation.
XX
KW Human; intracellular phosphorylation regulator; HRIP; stroke; myeloma;
KW neurological disorder; Parkinson's disease; demyelinating disease;
KW meningitis; developmental disorder; neuromuscular disorder; cancer;
KW myasthenia gravis; cell proliferative disorder; actinic keratosis;
KW arteriosclerosis; atherosclerosis; leukaemia; melanoma; bronchitis;
KW autoimmune disorder; inflammatory disorder; Addison's disease;
KW acquired immunodeficiency disease; allergy; diabetes mellitus;
KW rheumatoid arthritis; microbial infection; trauma; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 138..1505
FT FT /*tag= a
FT FT /product= "regulator of intracellular phosphorylation"
XX
XX WO200055332-A2.
XX
XX 21-SEP-2000.
XX
XX 17-MAR-2000; 2000WO-US07277.
XX
XX 18-MAR-1999; 99US-0125593.
XX
XX 20-MAY-1999; 99US-0135049.
XX
XX 09-JUL-1999; 99US-0143188.
XX
XX (INCYTE PHARM INC.
XX
XX Bandnan O, Tang YT, Yue H, Hillman JL, Baughn MR, Azimzai Y;
XX Lu DAM, Au-Young J;
XX
XX WPI; 2000-602121/57.
XX
XX P-PSDB; AAB18657.
XX
XX Novel human intracellular phosphorylation regulator polypeptides and
XX polynucleotides for diagnosis, prevention and treatment of
XX neurological, cell proliferative and autoimmune/inflammatory disorders
XX
XX Claim 4; Page 88; 96pp; English.
XX
XX The present sequence encodes a human regulator of intracellular
XX phosphorylation (HRIP). HRIP is useful for screening agonists and
XX antagonists of HRIP polypeptide. HRIP and its agonist or antagonist
XX are useful for treating a disease or condition associated with
XX decreased or increased expression of functional HRIP. Diseases treated
XX or diagnosed include neurological disorders such as stroke, Parkinson's
XX disease, demyelinating diseases, bacterial and viral meningitis and
XX other developmental disorders of the central nervous system,
XX neuromuscular disorders, myasthenia gravis, cell proliferative disorders
XX such as actinic keratosis, arteriosclerosis, atherosclerosis and cancer
XX including leukaemia, melanoma, myeloma and cancer of the adrenal gland,
XX bladder, bone, bone marrow, liver, lung, muscle, ovary, autoimmune/
XX inflammatory disorder such as Addison's disease, acquired
XX immunodeficiency diseases, allergies, bronchitis, diabetes mellitus,
XX rheumatoid arthritis, microbial infection and trauma.
```

```
XX
SQ Sequence 1706 BP; 509 A; 349 C; 427 G; 421 T; 0 other;
Alignment Scores:
Pred. No.: 6.3e-27 Length: 1706
Score: 482.00 Matches: 134
Percent Similarity: 48.75% Conservative: 81
Best Local Similarity: 30.39% Mismatches: 156
Query Match: 15.99% Indels: 70
DB: 21 Gaps: 18
US-09-830-144-2 (1-579) x AAA75674 (1-1706)
Qy 20 GluAlaProSerGlnValLeuAsnPheGluGluIleAspTyrLysGluLeuGluValGlu 39
Db 135 GAGATGCGTCTCTCGGTGCTCTTTGTGCAAAATAAATTTGATGACTTGCGAGTTTTT 194
Qy 40 GluValValGlyArgGlyAlaPheGlyValValCysLysAlaLysTrp-----ArgAla 57
Db 195 GAAAACTCGCGTGGAGGAAGTTTTGGGAGTGTATTATCGACCAATGGATATCACAGGAC 254
Qy 58 LysAspValAlaIleLys-----GlnIleGluSerGluSerGluArgLysAlaPhe 74
Db 255 AAGGAGTGGCTGTAAAGAAAGCTCTCAAAATAGAGAAAGAGGCAGAA----- 302
Qy 75 IleValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGly 94
Db 303 -----ATACTCAGTGTCTCTCAGTCACAGAAACATCATCCAGTTTTATGGA 347
Qy 95 AlaCysLeuAsnPro-----ValCysLeuValMetGluTyrAlaGluGlyGlySerLeu 112
Db 348 GTAATTTCTGAACCTCCCAACTATGCGATTGTGCAGAAATATGCTCTCTGGGATCCTC 407
Qy 113 TyrAsnValLeuHisGly-----AlaGluProLeuProTyrTyrThrAlaAlaHisAla 130
Db 408 TATGATTACATTAAACAGTAACAGAGTGAGGAGATG-----GATATGGATCATT 458
Qy 131 MetSerTrpCysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLys 150
Db 459 ATGACCTGGGCCACTGATGTAGCCAAAGGAATGCATTATTACATATGGAGGCTCTCTGTC 518
Qy 151 AlaLeuIleHisArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrVal 170
Db 519 AAGGTGATTACAGAGACCTCAAGTCAAGAAACGTTGTTAGCTGCTGATGGA---GTA 575
Qy 171 LeuLysIleCysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys 190
Db 576 TTGAAGATCTGTGACTTTGGT---GCCTCTCGGTTCCATAACCATACACACATGTCC 632
Qy 191 -----GlySerAlaAlaTrpMetAlaProGluValPheGluGlySerAsnTyrSerGlu 208
Db 633 TTGGTTGGAACCTTTCCCATGGATGGCTCCAGAAAGTTATCCAGAGTCTCCCTGTGTCAGAA 692
Qy 209 LysCysAspValPheSerTrpGlyIleLeuTrpGluValIleThrArgArgLysPro 228
Db 693 ACTGTGACACATATCTCTATGGTGTGTTCTCTGGGAGATGTAAACAGGAGGTCCTCC 752
Qy 229 PheAspGluIleGlyGlyProAlaPheArgIleMetTrp---AlaValHisAsnGlyThr 247
Db 753 TTTAAAGGTTTGGAAAGG-----TTACAAGTAGCTTGGCTTGTAGTGGAAAAAACGAG 806
Qy 248 ArgProProLeuIleLysAsnLeuProLysProIleGluSerLeuMetThrArgCysTrp 267
Db 807 AGATTAAACCAATCCAAAGCAGTTGCCCAAGAGTTTGTGTAACCTGTTTACATCAGTGTGG 866
Qy 268 SerLysAspProSerGlnArgProSerMetGluGluIleValLysIleMetThrHisLeu 287
Db 867 GAAGCTGATGCCAAAGACGCCCATCATCAAGCAAAATCATTTCAATCTCTG----- 917
Qy 288 MetArgTyrPheProGlyAlaAspGluProLeuGlnTyrProCysGlnTyrSerAspGlu 307
Db 918 -----GAG 920
```

Oy	308	GlyInSerAnSerAlaThrSerThrGlySerPheMetAspLeuIleAspThrAsnThr	327
Db	921	TCGATGCTAAATGACAGC-----AGCCTTCCTGCACAGTCACTCATTTCTTA	968
Oy	328	SerAsnIlySerAspThrAsnMetGluGlnValProAlaThrAsnAspThrIleLysArg	347
Db	969	CACAAACAAGCGCGAGTGGAGGTGGCGAA---ATTGAGGCACACTTGTGAGAGCTTAAAGAA	1025
Oy	348	LeuGluSerIlySleuLeuLysAsnGlnAlaLysGlnGlnSerGluSerIlyArgLeuSer	367
Db	1026	CTAGAGCCTGATCTTCAGCTTTAAGAGACAGACACTTAAAGAACGAGAAAGAGTTTAAAG	1085
Oy	368	Leu-----GlyAlaSerHisGlySerSerValGluSerLeuProProThrSer---	383
		::: :::	::: :::
Db	1086	ATGTGGAGGCAAAAAGCTGACAGACGACATCCAAACCCCGTCTCTTGGCTCTTCTGTCA	1145
Oy	384	-----GluGlyLysArgMetSerAlaAspMetSerGluIle	395
		-----	-----
Db	1146	AGAAATGCTGAGAGGTCTTACTTTGAATCTTAAACAGAGAGATCAAAAGCTGCACAGATG	1205
Oy	396	GluAlaArgGlyLeuAlaIleAlaThrThr---GlyAsnGlyGlnProArgArgSerIleGln	414
		::: :::	::: :::
Db	1206	TCATGTCTGCATCAACGACACAAAGTAACGSGGAGGCGCATGACCAACCAAGTGTGGAG	1265
Oy	415	AspLeuThrValIleThrglyThrGluProGlyGlnValSerSerArgSerSerSerProSer	434
		::: :::	::: :::
Db	1266	GCCATGATGCTGAGGCGCTTT-----GGGATATCTTCTCAATGAAACAAAGCAGAGACT	1319
Oy	435	Val 435	
Db	1320	G TG 1322	

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 9, 2002, 22:59:44 ; Search time 1808.11 Seconds
(without alignments)
5186.195 Million cell updates/sec

Title: US-09-830-144-2
Perfect score: 3014
Sequence: 1 MSTAASSSSSSSAGEMIE.....QCKQLEIVRSQQQKROQTS 579

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2.1/USPTO.spool/US09830144/runat_04122002.141353.2264/app.query.fasta_1.1422
-DB=EST_QFMT=faststep -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPELU=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09830144 @cgn.1.1.2441 @runat_04122002.141353.2264 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOF=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: em_estba:*
2: em_esthum:*
3: em_estcin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_estcl:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_fod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1443.5	47.9	956	14	BQ958174	BQ958174 AGENCOURT
2	1379	45.8	1062	13	BM554120	BM554120 AGENCOURT
3	1242.5	41.2	998	9	AL525728	AL525728 AL525728
4	1225	40.6	771	13	BI093821	BI093821 602860601
5	1212	40.2	739	13	BI696710	BI696710 603347214
6	1140	37.8	768	13	BI868773	BI868773 603391922
7	1119	37.1	827	9	AL550589	AL550589 AL550589
8	1116	37.0	688	13	BQ062988	BQ062988 BQ062988
9	1116	37.0	696	13	BJ074867	BJ074867 BJ074867
10	1112	37.0	1006	14	BQ219348	BQ219348 AGENCOURT
11	1112	36.9	663	10	BB617944	BB617944 BB617944
12	1086	36.0	646	13	BJ073883	BJ073883 BJ073883
13	1079	35.8	604	13	BM426610	BM426610 P912n.pk0
14	1063	35.3	929	9	AL520975	AL520975 AL520975
15	1061	35.2	686	12	BG548917	BG548917 602575750
16	1055.5	35.0	868	12	BE742348	BE742348 601575585
17	1047	34.7	858	13	BG916821	BG916821 602816079
18	1009.5	33.5	910	12	BF780358	BF780358 602103276
19	1006	33.4	641	14	BQ386875	BQ386875 NISC.mn20
20	995	33.0	621	10	BB618631	BB618631 BB618631
21	994.5	33.0	1054	14	BQ067693	BQ067693 AGENCOURT
22	992	32.9	711	10	AV718168	AV718168 AV718168
23	955	31.7	968	12	BG027707	BG027707 602294752
24	954	31.7	599	10	BE284386	BE284386 601087550
25	943.5	31.3	661	12	BF678411	BF678411 602085846
26	923.5	30.6	941	12	BG029948	BG029948 602297101
27	923	30.6	551	9	AA461278	AA461278 2x63x04.Y
28	921	30.6	747	13	BI334774	BI334774 602999677
29	920	30.5	542	10	AW960377	AW960377 EST372448
30	912	30.3	822	13	BI757369	BI757369 603029591
31	901.5	29.9	586	13	BQ035129	BQ035129 BQ035129
32	880.5	29.2	760	12	BG773488	BG773488 602720209
33	873	29.0	682	10	AV398933	AV398933 AV398933
34	866	28.7	602	10	BE080188	BE080188 RC4-BT062
35	852.5	28.3	745	13	BI767913	BI767913 603061057
36	843	28.0	681	10	AV398935	AV398935 AV398935
37	827.5	27.5	1191	13	BM472107	BM472107 AGENCOURT
38	814	27.0	1033	12	BE882182	BE882182 601505283
39	811.5	26.9	673	9	AJ392980	AJ392980 AJ392980
40	809	26.8	647	10	BB665645	BB665645 BB665645
41	794.5	26.4	594	13	BJ014293	BJ014293 BJ014293
42	782	25.9	729	12	BG826184	BG826184 602750551
43	750	24.9	671	12	BG700279	BG700279 602679903
44	741.5	24.6	585	9	AL773651	AL773651 AL773651
45	740.5	24.6	718	9	AL652289	AL652289 AL652289

ALIGNMENTS

RESULT 1
BQ958174
LOCUS
DEFINITION AGENCOURT_10014737 NIH_MGC_101 Homo sapiens cdna clone
IMAGE:6497311 5', mRNA sequence.
ACCESSION BQ958174
VERSION BQ958174.1 GI:22373652
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 956)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LLM2678 row: c column: 08
High quality sequence stop: 686.
Location/Qualifiers

FEATURES

source

1..956
/organism="Homo sapiens"
/db xref="taxon:9606"
/clone="IMAGE:6497311"
/clone_lib="NIH_MGC_101"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; Site 1: EcoRI; Site 2:
XhoI; CDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

BASE COUNT 330 a 222 c 199 g 204 t 1 others
ORIGIN

Alignment Scores:

Pred. No.: 5,97e-126 Length: 956
Score: 1443.50 Matches: 290
Percent Similarity: 95.10% Conservative: 1
Best Local Similarity: 94.77% Mismatches: 11
Query Match: 47.89% Indels: 4
DB: 14 Gaps: 2

US-09-830-144-2 (1-579) X BQ958174 (1-956)

QY 263 MetThArgCysTrpSerIysAspProSerGlnArgProSerMetGlnGluIleValLys 282
DB 2 ATGACTCGTGTGGTCTTAAGATCCTTCCACGGCCCTTCAATGAGAAATGTGAAA 61
QY 283 IleMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyrProCys 302
DB 62 ATTAATGACTCACTTGATCGCGTACTTCCAGGAGCAGATGAGCATTACGATACCTTGT 121
QY 303 GlnTyrSerAspGluGlnSerAsnSerAlaThrSerThiGlySerPheMetAspIle 322
DB 122 CAGATTCAGATGAGAGACAGACCAACTGCGCACAGATACAGGCTCATTCATGACATT 181
QY 323 AlaSerThrAsnThrSerAsnLysSerAspThrAsnMetGlnGlnValProAlaThrAsn 342
DB 182 GCTTCTCAAAATCGAGTAACAAAGGACACTTAATATGAGCAAGTCCGCGCAAAAT 241
QY 343 AspThrIleLysArgLeuGlnSerLysLeuLeuLysAsnGlnAlaLysGlnGlnSerGlu 362
DB 242 GATACCTTTAAGCGCTTAGAATCAAATTTCTTAAAAATCAGGCACAAAGACAGATGAA 301
QY 363 SerGlyArgLeuSerLeuGlyAlaSerHisGlySerSerValGluSerLeuProPheThr 382
DB 302 TCTGAGAGTTTAAGCTTGGAGCGCTCCCGTGGAGACAGTGTGAGAGCTTGGCCCCAAC 361
QY 383 SerGlnGlyLysArgMetSerAlaAspMetSerGlnIleGlnAlaArgIleAlaIaThr 402
DB 362 TCTGAGGCAAGAGAGATGAGATGCTGACATGCTGAATAATAGACCTTAGATCGCCGAA 421
QY 403 ThrGlyAsnGlyGlnProAlaArgArgSerIleGlnAspLeuThrValThrGlyThrGln 422
DB 422 ACGAGCAACGACAGCAACGACGTAAGATTCATCAAGACTTGAAGCTGTAAGTGAACAGAA 481

QY 423 ProGlnValSerSerArgSerSerSerProSerValArgMetIleThrThrSerGly 442
DB 482 CTTGTCAGATGAGACAGTATGATCATCCAGTCCAGATGATATGATCTACCTGAGA 541
QY 443 ProThrSerGlnLysProThrArgSerHisProTyrThrProAspAspSerThrAspThr 462
DB 542 CCAACCTCCAAAAGCCAACTCCAAAGTCAATGATCCATGAGACCCCTGATATTCACAGATAC 601
QY 463 AsnGlySerAspAsnSerIleProMetAlaTyrLeuThrLeuAspHisGlnLeuGlnPro 482
DB 602 AATGATCAGATTAATCTCATCCCAATGGCTTATCTTACACTGGATCACAACATCAGCCT 661
QY 483 LeuAlaProCysProAsnSerLysGluSerMetAlaValPheGlnGlnHisCysLysMet 502
DB 662 CTACACCGGCGCCCAAACTCCAAAGATCTATGACAGTGTGAAAGCATGTGTAATG 721
QY 503 AlaGlnGlnTyrMetLysValGlnThrGlnIleAlaLeuLeuLeuGlnArgLysGlnGlu 522
DB 722 GCACAAATATTTATGAAAGTTCAAACGAAATGCAATGTTATTTACAGAGAAAGCAAGAA 781
QY 523 LeuValAlaGluLeuAspGlnAspGlnLysAspGlnGlnAsnThrSerArgLeuVal-GI 542
DB 782 CTAGTTCAGAACTGACGACGACGATGAAAGGAGACAGCAAAATACATCTCGCTGATAC 841
QY 542 nglnHisLysLysLeuLeu---AspGluAsnLysSerLeu-----SerThrTyrTyrGI 559
DB 842 GGAACATTAATAAAAGCTTTTAAATGAAAACAAAGCCTTTCTACCTTACCTACAGNC 901
QY 559 nglnCysLysLysGln 564
DB 902 AATGCCAAAATAAAACA 917

RESULT 2
BM554120 1062 bp mRNA linear EST 20-FEB-2002
AGENCY/URL 6546903 NIH_MGC_119 Homo sapiens CDNA clone IMAGE:5742353
DEFINITION 5', mRNA sequence.
ACCESSION BM554120
VERSION BM554120.1 GI:18793446
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1062)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LLM2760 row: f column: 18
High quality sequence stop: 730.
Location/Qualifiers

FEATURES

source

1..1062
/organism="Homo sapiens"
/db xref="taxon:9606"
/clone="IMAGE:5742353"
/clone_lib="NIH_MGC_119"
/tissue_type="medulla"
/lab_host="DH10B"
/note="Organ: brain; Vector: PCWV-SPORT6; Site 1: NotI;
Site 2: EcoRV (destroyed); RNA source normal medulla from
anonymous male aged 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range
0.9-3 kb. Library is normalized and enriched for

Db 429 ATTATCCCGGTGACACCTCTAATATGTAAGCTTTATGAGCCTGCTGAATCCACT 488
Qy 100 |CysLeuValMetGluTyr-AlaGluGlyGlySerLeuTyrAsnValLeuHisGlyAlaG 120
Db 489 GTGCTCTTGATGGAATATAGCTGAAGGGGGCTTTATATATGATGCTGACGTGCTG 548
Qy 120 |LurLeuLeuProTyrTyrThAlaAlaHisAlaMetSerTPrpCysLeuGlnCysSerGlnG 140
Db 549 AACCATTTGCCATATTAATACGCTGCCACGCAATGARTTGTTGTTACAGTGTTCACAG 608
Qy 140 |ValAlaTyrLeuHisSerMetGlnProValAlaLeuLeuHisArgAspLeuLysProp 160
Db 609 GAGTGGCTTATCTTCACAGATCCAAACCGCTAATTCACAGGACCTGAAACAC 668
Qy 160 roAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThAlaC 180
Db 669 CAACCTACTGCTGTTGACGAGGGGACAGTCTTAAAAATTTGTGATTTGGTACAGCCT 728
Qy 180 yAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaATPMeAlaProGluV 200
Db 729 GTACATTCAGACACATGACCAATTAACAAGGAGTGTCTGCTTGATGGACCTGAAG 788
Qy 200 alPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTPrpGlyIleIleLeuT 220
Db 789 TTTTGAAGGTAGTATTAAGTGAAGAAAATGTACGCTCTTCAGCTGGGTATTTATCTTT 848
Qy 220 rpgIuValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleM 240
Db 849 GGGAAGATGATMG-CGTGGAAACCTTT-GATGAGATGTGTGCCAGCTTCGGAATCA 906
Qy 240 ecTPrAlaVal-HisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIle 259
Db 907 TGTGGGCTGTCTTCAATATGTAAGTCTGACCACTGATTAATAATTACCTAAGCATT- 965
Qy 260 GluSerLeuMetThrArgCysTPrSerLys 269
Db 966 GAGAGCCTGATGACCTGCTTGTGTTTAAA 995

RESULT 4
LOCUS B1093821 771 bp mRNA linear EST 20-JUN-2001
DEFINITION 602860601F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:5001504 5',
mRNA sequence.
ACCESSION B1093821
VERSION B1093821.1 GI:14512151
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 771)
AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
plate: LLAM1035 row: b column: 01
High quality sequence stop: 768.

FEATURES
Location/Qualifiers
1..771

source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5001504"
/clone_1id="NIH_MGC_10"
/cell_line="MG36"
/lab_host="DH10B"

/note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5 kb. Library prepared by Life
Technologies."
BASE COUNT 230 a 163 c 187 g 191 t
ORIGIN
Alignment Scores:
Pred. No.: 1,96e-105 Length: 771
Score: 1225.00 Matches: 232
Percent Similarity: 97.92% Conservative: 3
Best Local Similarity: 96.67% Mismatches: 5
Query Match: 40.64% Indels: 0
DB: 13 Gaps: 0

US-09-830-144-2 (1-579) x B1093821 (1-771)

Qy 171 LeuLysIleCysAspPheGlyThAlaCysAspIleGlnThrHisMetThrAsnAsnLys 190
Db 1 CTTAAAAATTTGTGATTTTGTGACAGCTGTGACCTTGACACCTTGACACACATGACCAATACAG 60
Qy 191 GlySerAlaAlaATPMeAlaProGluValPheGluGlySerAsnTyrSerGluLysCys 210
Db 61 GGAGTGTGCTGTGATGATGGACACCTGAAAGTTTGAAGGTATTAATTAACGTGAATAATGT 120
Qy 211 AspValPheSerTPrpGlyIleIleLeuTPrpGluValIleThrArgArgLysProPheAsp 230
Db 121 GACGCTTCACAGCTGGGGATATATCTTTGGGAAGTGTATTAACGCTCGGAACCCCTTTGAT 180
Qy 231 GluIleGlyGlyProAlaPheArgIleMetTrpAlaValHisAsnGlyTPrArgProPro 250
Db 181 GAGATGTGTGCCAGCTTCGGAATCATGTGGCTGTTCATATATGTAAGTCTGACACCA 240
Qy 251 LeuIleLysAsnLeuProLysProIleGluSerLeuMetThrArgCysTPrSerLysAsp 270
Db 241 CTGATTAATAAATTTTAACTTAAGCCATTGAGAGCTGATGATGATGCTGTTGTGTAAAGAT 300
Qy 271 ProSerGlnArgProSerMetGluGluIleValLysIleMetThrHisLysMetArgTyr 290
Db 301 CCTTCCCGACCGCCCTTCATATGAGGAATGTGAAAATTAATGACTCACTGATGCGGTAC 360
Qy 291 PheProGluValAspGluProLeuGlnTyrProCysGlnTyrSerAspGluGlnSer 310
Db 361 TTCCAGAGCAGATGAGCCATTACAGATCTTGTCAGATTCAGATGAGGAAGACAGAGC 420
Qy 311 AsnSerAlaThrSerThrGlySerPheMetAspIleAlaSerThrAsnThrSerAsnLys 330
Db 421 AACTCTGCCACACATGACAGGCTCATTCATGACATTGCTTCAAAATTAAGATGACAA 480
Qy 331 SerAspThrAsnMetGluGlnValProAlaThrAsnAspThrIleLysArgLeuGlnSer 350
Db 481 AGTGACACTAATATATGAGCAAGTCTCGCACCAATATATCTAATTAAGCCCTTGAATCA 540
Qy 351 LysLeuLeuLysAsnGlnAlaLysGlnGlnSerGlnSerGlyArgLeuSerLeuGlyAla 370
Db 541 AAATCGTTGAAAATTAACGGCAACCAACAGAGTGAATCTGACCTTTAACTTGGGAGCC 600
Qy 371 SerHisGlySerSerValGluSerLeuProProThrSerGlnGlyLysArgMetSerAla 390
Db 601 TCCCTGGGAGACAGTGTGAGAGCTTGCCCCCAAGCTCTGAAGGCAAGAGATAGTGCT 660
Qy 391 AspMetSerGluIleGluAlaArgIleAlaAlaThrThrGlyAsnGlyGlnProArgArg 410
Db 661 GACATGTCTGAATATGAAGATGATGCGCGCAACCAACAGCCTATTCCAAAGCTTAAACGG 720

RESULT 5
B1696710

LOCUS B1696710 739 bp mRNA linear EST 18-SEP-2001
DEFINITION 603347214F1 NCI_CGAP_Mam2 Mus musculus cDNA clone IMAGE:5374971 5',
mRNA sequence.
ACCESSION B1696710
VERSION B1696710.1 GI:15659339
KEYWORDS EST.

SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M11954 row: k column: 04
High quality sequence stop: 717.
FEATURES
source
1..739
/organism="Mus musculus"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone="IMAGE:5374971"
/clone_lib="NCI_CGAP_Mam2"
/tissue_type="tumor, biopsy sample"
/dev_stages="5 months"
/lab_host="DH10B"
/note="organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
Providing samples: Gilbert Smith, NIH"
BASE COUNT 195 a 169 c 188 g 187 t
ORIGIN
Alignment Scores:
Pred. No.: 3,1e-104 Length: 739
Score: 1212.00 Matches: 235
Percent Similarity: 97.93% Conservative: 2
Best Local Similarity: 97.11% Mismatches: 5
Query Match: 40.21% Indels: 0
DB: 13 Gaps: 0
US-09-830-144-2 (1-579) x BI696710 (1-739)
QY 83 ArgValAsnHisProAsnLeValIysLeuTyrGlyAlaCysLeuAsnProValCysLeu 102
Db 3 CGTGTGAACCATCTTAACATTTGTCAGGAGCTGCTGAATCCAGTATGTCTT 62
QY 103 ValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGlyAlaGluProLeu 122
Db 63 GTGATGGAATATGCAGAGGGGGCTCAITTTGTAATGTCTGTCATGGTCTGAACCATTTG 122
QY 123 ProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeuGlnCysSerGlnGly-ValAl 142
Db 123 CCTTACTACACTCTGCTCATGCATGAGTGTGTGTTCAGTGTTCACAGGTAGTGGC 182
QY 142 aTyrLeuHisSerMetGlnProIlysalalaLeuIleHisArgAspLeuIysProAsnLe 162
Db 183 TTACCTGCACAGCATGCAGCCCAAGCGCTGATTTCACAGGAGCTCAAGCTCCAAACTT 242
QY 162 uLeuLeuValAlaGlyGlyThrValLeuIlysalalaCysAspPheGlyThrAlaCysAsp 182
Db 243 GCTGCTGGTTGCAGAGGGGACATTTCTAAATAATCTCGATTTTGGTACAGCTTGTGACAT 302
QY 182 eGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyrMetAlaProGluValPheG 202
Db 303 CCAACACACATGACCCATTAATAAGGAGTGTGCTTGGATGGCCCTGAAGTGTTCGA 362
QY 202 uGlySerAsnTyrSerGluIysCysAspValPheSerTyrGlyIleIleLeuTyrPLeuVa 222
Db 363 AGGTAGCAATTACAGTGAAAGGTGTGATGCTTCAGCTGGGGTATTATCTCTCTGGGAAGT 422

QY 222 lIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMetTyrAl 242
Db 423 GATAACACCCCGAAACCTTCATGATGATCGTGGCCAGCTTTCAGAAATCATGTGGC 482
QY 242 aValHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGluSerLe 262
Db 483 TGTTTATAATGGCACTCGACCACTGATCAAAAATTTACCTAAGCCCATTTGAGAGCTT 542
QY 262 uMetThrArgCys-TyrSerLysAspProSerGlnArgProSerMetGluGluIleValL 282
Db 543 GATGACACACCTGTGTCTTAAGGACCCATCTCAGGC-CTTCAATGGAGAAATTTGA 601
QY 282 ySileMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyrProC 302
Db 602 AAATAATGACTACTTGTATCGGTACTTCCAGGAGCGATGAGCCATTACAGTATCCTT 661
QY 302 ySgln-TyrSerAspGluGlyGlnSerAsnSerAlaThrSerThrGlySerPheMetAsp 321
Db 662 GTCAGGTACTCTGATGAAGGGCAGAGCAATCC-AGCACCAGCAGGCTCGTTTCATGAA 720
RESULT 6
BI686773 768 bp mRNA linear EST 11-OCT-2001
LOCUS 603391922F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:5402236 5',
DEFINITION mRNA sequence.
ACCESSION BI686773
VERSION BI686773.1 GI:16042446
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 768)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M12025 row: k column: 05
High quality sequence start: 2
High quality sequence stop: 761.
FEATURES
source
1..768
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5402236"
/clone_lib="NIH_MGC_90"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH MGC Library."
BASE COUNT 271 a 177 c 164 g 156 t
ORIGIN
Alignment Scores:
Pred. No.: 2,14e-97 Length: 768
Score: 1140.00 Matches: 243
Percent Similarity: 95.72% Conservative: 3
Best Local Similarity: 94.55% Mismatches: 10
Query Match: 37.82% Indels: 5
DB: 13 Gaps: 0

US-09-830-144-2 (1-579) x B1668773 (1-768)

QY 315 SerThrGlySerPheMetAspIleAlaSerThrAsnThrSerAsnLysSerAspThrAsn 334
 DB 3 AGTACAGGCGCTCATGACATGCTTCTTACAAATACAGATACAAAGTGACACTAAT 62
 QY 335 MetGluGlnValProIleThrAsnAspThrIleValArgLeuGlnLysSerLeuLys 354
 DB 63 ATGGACCAAGTCTCTCCCAAAATGATTAATTAAGCGCTTAGAAATCAAAATG-TTGAAA 121
 QY 355 AsnGlnAlaLysGlnInserGlnSerGlyArgLeuSerLeuGlyAlaSerHisGlySer 374
 DB 122 AATCAGCAAGCAAGCAAGATGTAATCTGACGTTTAAGTTTGGAGCCCTCCGTGGAGC 181
 QY 375 SerValGluSerLeuProProThrSerGlnGlyLysArgMetSerAlaAspMetSerGlu 394
 DB 182 AGGTGGAAGAGCTTGCCCCCAACCTGTGAGGCAAGAGATGATGCTGACATGCTCGAA 241
 QY 395 ILeGluAlaArgIleAlaIleThrThrGlyAsnGlyGlnProArgArgSerIleGln 414
 DB 242 ATGGAAGCTAGAGTCCGCCCAACACAGGCAACGACACGAAAGCTAGATCCATCCAA 301
 QY 415 AspLeuThrValThrGlyThrGluProGlyGlnValSerSerArgSerSerProSer 434
 DB 302 GACTTGACTGTAAGTGAACAGAACTGCTCAGGTGAGCAGTATGATCCAGTCCAGT 361
 QY 435 ValArgMetIleThrThrSerGlyProThrSerGlyLysProThrArgSerHisProThr 454
 DB 362 GTCAGATGATTAATTAATCTCAGGACCAACCTCAGAAAGCCAACTGAAATCATCATG 421
 QY 455 ThrProAspAspSerThrAspThrAsnGlySerAspAspSerIleProMetAlaIleLeu 474
 DB 422 ACCCTGATGATTCACAGATACCAATGATCAATTAATCCATCCCAATGGCTTATCTT 481
 QY 475 ThrLeuAspHisGlnLeuGlnProLeuAlaProCysProAsnSerLysGlnSerMetAla 494
 DB 482 ACACGTGATCAACCACTACAGCCTCTAGACCGTGCCCAAACTCCAAAGATCTATGCA 541
 QY 495 ValPheGlnGlnHisCysLysMetAlaGlnGlyIleMetLysValGlnThrGluIleAla 514
 DB 542 GGTGTTGAACACGATGTAATAATGGCACAAATATATGAAGTTCAAAACGAATATGAT 601
 QY 515 LeuLeuLeuGlnArgLysGlnLeuValAla-GluLeuAspGlnAspGluLysAspGln 534
 DB 602 CGCTTATTAACAGAAACCAAACTAGTTGCAAGACTGACCGAGATGAAAAGGACCA 661
 QY 534 GlnAsnThrSerArgLeuValGlnGlnHisLysLysLeuLeuAspGlnAsnLysSerLe 554
 DB 662 GCAAAATATCATCTCGCTGTGACGAGAACATTAAGCTTTAGATG-AAAACAAAAGACT 720
 QY 554 userthrtyrtyrtyrtyrtyrtyrtyrtyrtyrtyrtyrtyrtyrtyrtyrtyrtyr 570
 DB 721 TTTCTACTTAACACAGCA-TGCMAAAACA-CTAGAGGTCTATCAGAACT 767
 RESULT 7
 AL550589 827 bp mRNA linear EST 16-FEB-2001
 LOCUS AL550589 LTI_NFL006.PL2 Homo sapiens cDNA clone CS01058Y014.5
 DEFINITION prime, mRNA sequence.
 ACCESSION AL550589
 VERSION AL550589.1 GI:12887700
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 827)
 AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France

Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr,
 Location/Qualifiers
 1. 827
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS01058Y014"
 /clone_lib="LTI_NFL006.PL2"
 /tissue_type="placenta"
 /note="Vector: pCWS/PORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT)-primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCWS/PORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
 Email : fliang@life.com URL : <http://fulllength.invitrogen.com>

BASE COUNT 196 a 178 c 228 g 209 t 16 others
 ORIGIN

Alignment Scores:
 Pred. No.: 2,33e-95 Length: 827
 Score: 1119.00 Matches: 216
 Percent Similarity: 97.32% Conservative: 2
 Best Local Similarity: 96.43% Mismatches: 5
 Query Match: 37.13% Indels: 1
 DB: 9 Gaps: 0

US-09-830-144-2 (1-579) x AL550589 (1-827)

QY 1 MetSerThrAlaSerAlaAspSerSerSerSerSerAlaGlyLysMetIleGln 20
 DB 157 ATGCTTAMAGCCTCTGCGGCTCTCMTCTCTGCTTCCGCGGTARATGATCGAA 216
 QY 21 AlaProSerGlnValLeuAsnPheGlnGluIleAspTyrLysGluIleGluValGluGlu 40
 DB 217 GCCCTTCCAGGTCCTCACTTTGAAGATCGATCAACAGAGATCCAGGTGGAAAG 276
 QY 41 ValValGlyArgGlyValaPheGlyValValCysLysAlaLysTyrArgAlaLysAspVal 60
 DB 277 GTTGTTGGAAGAGAGACCTTTGGARTTTTGCAAGCTAGATGGAGACCAAAATGTT 336
 QY 61 AlaIleLysGlnIleGlnSerGlnSerGlnArgLysAlaPheIleValGluLeuArgGln 80
 DB 337 GCTATTAAATAATAAAGTGAATCTGAGAGAAAGCTTTATTGTAGACTTGGGAG 396
 QY 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnProVal 100
 DB 397 TTATCCCGTGAACCATCTTAATATTGTAAAGCTTTATGAGGCTGTGTAATCCAGTG 456
 QY 101 CysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGlyAlaGlu 120
 DB 457 TGTCTTGATGGAATATCTGAAGGGGCTCTTTATATATATGTCATGATGCTGTA 516
 QY 121 ProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeuGlnCysSerGln-Gl 140
 DB 517 CCATTCCCATATTATATCTCTGCCACGCAATGAGATTGGTGTTCACAGTCCCAAG 576
 QY 140 yValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysProPr 160
 DB 577 AGTGGCTTATCTTCAAGGATGCAACCCAAAGCGCTTAATTCACAGGAGCTGAAACAC 636
 QY 160 oAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCys 180
 DB 637 AAACCTACAGCTGTTGGACGGGGAGACGTTTAAATTTTGATTTTGGTACGCTG 696
 QY 180 sAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyrMetAlaProGluVal 200
 DB 697 TGACATTCAAGACACATGACCAATTAACAAGGGAGTGTCTGTTGATGACCACTGAAGT 756
 QY 200 lPheGluGlySerAsnTyrSerGluLysCysAspValaPheSerTyrGlyIleIleLeuTr 220

```

Db 757 TTTTGAAGGTAGTAATTACAGTCGAAAAAATGTGACGTCTTCAGCTGGGTATTATTCTTTG 816
Qy 220 pgluValile 223
Db 817 GGAAGTGATA 826

RESULT 8
BJ062988 688 bp mRNA linear EST 10-DEC-2001
LOCUS BJ062988
DEFINITION laevis cDNA clone XL069m17 5', mRNA sequence.
ACCESSION BJ062988
VERSION BJ062988.1 GI:17470746
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 688)
AUTHORS Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara
,Y.
TITLE Expressed genes in X. laevis embryo
JOURNAL Unpublished (2001)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
source
1..688
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="XL069m17"
/library="NIBB Mochii normalized Xenopus tailbud
library"
/tissue_type="whole embryo"
/dev_stage="stage 25"
/notes="Vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; cDNAs
were oligo-dr primed and directionally cloned. Staging
according to Nieuwkoop and Faber. Library is subtracted
and was constructed by N. Garrett and A.M. Zorn,
(Wellcome/CRC Institute)."
```

BASE COUNT 206 a 139 c 154 g 188 t 1 others

ORIGIN

Alignment Scores:

Pred. No.:	3.33e-95	Length:	688
Score:	1116.00	Matches:	204
Percent Similarity:	96.31%	Conservative:	5
Best Local Similarity:	94.01%	Mismatches:	8
Query Match:	37.03%	Indels:	0
DB:	13	Gaps:	0

US-09-830-144-2 (1-579) x BJ062988 (1-688)

```

Qy 120 GluProLeuProTyrThrAlaAlaHisAlaMetSerTrpCysLeuGlnCysSerGln 139
Db 22 GAACCTTTGCTTACTATCTACTGCTGCCATGCAATGAGTTGGTGTTCACATGTGCCCAA 81
Qy 140 GlyValAlaTyrLeuHisSerMetGlnProLeuAlaLeuIleHisArgAspLeuLysPro 159
Db 82 GGAGTTGCATATTTACATGATGAGCCAAAGGCTCTGATTACAGGACCTCAAAACCA 141
Qy 160 ProAsnLeuLeuLeuValAlaGlyThrValLeuLysIleCysAspPheGlyThrAla 179
Db 142 CCAAACTGTTGTGTAGCTGGAGCAGCTGTTCTTAAGATTGTGACTTTGGTACAGCC 201
Qy 180 CysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGlu 199
Db 202 TGTGATATTACAGCTCACATGACTAATAACAAAGGAAGTGCAGCATGGATGGCTTCAGAA 261
```

```

Qy 200 ValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrpGlyIleIleIeu 219
Db 262 GTTTTGAAGGTAGCACTACAGCGAAAAAATGTGACGTGTTAGTTGGGGCATTTATTTCTT 321
Qy 220 TrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgile 239
Db 322 TGGGAAGTAATAACCGGAAGAAACCTTTTCGATGAAATTTGGTGGTCCAGCGTTCCGTATA 381
Qy 240 MetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIle 259
Db 382 ATGTGGGCTGTTCAATGTGCTACTCGGCCACCATTAATAATAAAATTTGCCCTAACCTATT 441
Qy 260 GluSerLeuMetThrArgCysTyrSerLysAspProSerGlnArgProSerMetGluGlu 279
Db 442 GAAAGCTTAATGACTCGCTGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 501
Qy 280 IleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGln 299
Db 502 ATTGTCAAGATAATGACACATCTAAAGCAGTATTTTCTCGAGCAGACGTTTCCTTACAG 561
Qy 300 TyrProCysGlnTyrSerAspGluGlyGlnSerAsnSerAlaThrSerThrGlySerPhe 319
Db 562 TATCCTTTGTCAGTACTCTGATGAAAGGCAAGCAATTTCTGCAACAGTACAGGTTCTCTGT 621
Qy 320 MetAspIleAlaSerThrAsnThrSerAsnLysSerAspThrAsnMetGlu 336
Db 622 ACTGATATTACATCTACTAATACTAGTAACAAGAGTGAGTCAATATAGAA 672

RESULT 9
BJ074867 696 bp mRNA linear EST 11-DEC-2001
LOCUS BJ074867
DEFINITION laevis cDNA clone XL071110 5', mRNA sequence.
ACCESSION BJ074867
VERSION BJ074867.1 GI:17505056
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 696)
AUTHORS Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara
,Y.
TITLE Expressed genes in X. laevis embryo
JOURNAL Unpublished (2001)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
source
1..696
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="XL071110"
/library="NIBB Mochii normalized Xenopus tailbud
library"
/tissue_type="whole embryo"
/dev_stage="stage 25"
/notes="Vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; cDNAs
were oligo-dr primed and directionally cloned. Staging
according to Nieuwkoop and Faber. Library is subtracted
and was constructed by N. Garrett and A.M. Zorn,
(Wellcome/CRC Institute)."
```

BASE COUNT 210 a 142 c 155 g 188 t 1 others

ORIGIN

Alignment Scores:

Pred. No.:	3.39e-95	Length:	696
------------	----------	---------	-----

Score: 1116.00 Matches: 204
Percent Similarity: 96.31% Conservative: 5
Best Local Similarity: 94.01% Mismatches: 8
Query Match: 37.03% Indels: 0
Gaps: 13

US-09-830-144-2 (1-579) x BJ074867 (1-696)

QY 120 GluProLeuBroTyTyrThrAlaAlaHisAlaMetSerTrpCysLeuGlnCysSerGln 139
DB 22 GAACCTTGGCTTACTACTGCTGCTGCCCATGATGATGGTGTTCATCAATGTGCCAA 81
QY 140 GlyValAlaIleuHisSerMetGlnProIleValLeuIleHisArgAspLeuPro 159
DB 82 GGAGTTGCAATTTTACATGACATGAAGCCAAAGGCTGTGATTCACAGGAGCCCAACCA 141
QY 160 ProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAla 179
DB 142 CCAAACTTGTCTGCTGAGCTGAGGACACTGTTCTTAAGATTGTGACTTGGTACAGCC 201
QY 180 CysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGlu 199
DB 202 TGTGATTTTCAGATCAGATCAGATTAATTAACAAGAAAGTGCAGATGATGGCTCCAGAA 261
QY 200 ValPheGluGlySerAsnTrpSerGluLysCysAspValPheSerTrpGlyIleIleLeu 219
DB 262 GTTTTGAAGGTAGCACTACAGCAAAATGTGACGTGTTTGTGGGGCATTTATCTT 321
QY 220 ThrGluValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIle 239
DB 322 TGGGAATTAATTAACCCGAAAGAAACCTTTGATGAATTTGGTGTCCAGGCTTCCGTA 381
QY 240 MetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIle 259
DB 382 ATGTGGCTGTTCACATGATGATCGGCGCACATTAATTAATAATTGGCTTAAGCTAT 441
QY 260 GluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgProSerMetGluGlu 279
DB 442 GAAAGCTTAATGACTCGCTCTGTGTCMAAGATCCCCCAAAAGACCTTCAATGAGGAG 501
QY 280 IleValLysIleMetThrHisIleMetArgTrpPheProGluValAspGluProLeuGln 299
DB 502 ATGTCAAGTAATGACACATCTTAAGACATATTTCTTGAGACACGCTTCTTAACG 561
QY 300 TyrProCysGlnTrpSerAspGluGlnSerAsnSerAlaThrSerThrGlySerPhe 319
DB 562 TATCTTGTAGTACTGTGATGAAGGCAAGCATTTCTGCAACAAGTACAGGTTCTGT 621
QY 320 MetAspIleAlaSerThrAsnThrSerAsnLysSerAspThrAsnMetGlu 336
DB 622 ACTGATATTACATCTACTAATACTAGTACAGTACAGTCAATATAGAA 672
RESULT 10
LOCUS BQ219348 1006 bp mRNA linear EST 02-MAY-2002
DEFINITION AGENCOURT_7260864 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5785294
5', mRNA sequence.
ACCESSION BQ219348
VERSION BQ219348.1 GI:20400748
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12872 row: c column: 23
High quality sequence stop: 624.

FEATURES

SOURCE

1..1006
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5785294"
/clone_lib="NIH MGC 71"
/tissue_type="leiomysarcoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site: 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.1 kb."
BASE COUNT 240 a 240 c 293 g 233 t
ORIGIN

Alignment Scores:

Pred. No.: 6,156-95 Length: 1006
Score: 1116.00 Matches: 226
Percent Similarity: 84.31% Conservative: 5
Best Local Similarity: 82.48% Mismatches: 17
Query Match: 37.03% Indels: 27
Gaps: 4

US-09-830-144-2 (1-579) x BQ219348 (1-1006)

QY 1 MetSerThrAlaSerAlaAlaSerSerSerSerSerSerAlaGlyGluMetIleGlu 20
DB 163 ATGTCTAAGGCTGTGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 222
QY 21 AlaProSerGlnValIleAsnPheGluGluIleAspTrpLysGluIleGluVal 40
DB 223 GCCCTTCCCGAGTCTCTCACTTTGAAGAGATCAACAGAGATCGAGTGGAGAG 282
QY 41 ValValGlyArgGlyAlaPheGlyValValCysLysAlaLysTrpArgAlaLysAspVal 60
DB 283 GTTGTGGAAGAGAGAGCTTTGAGTTGTTGCAAACTCAAGTGAAGCAAAAGATGT 342
QY 61 AlaIleLysGlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
DB 343 GCTATTAAACAAATGAAGATGAGTGAAGGAAACGTTTATTTAGAGCTTCGGCAG 402
QY 81 LeuSerArgValAsnHisProAsnIleValLysLeuTrpGlyAlaCysLeuAsnProVal 100
DB 403 TTATCCCGTGGAAACCATCTTAATTTGTAAAGCTTTATGAGCGCTGATCCAGTG 462
QY 101 CysLeuValMetGluTrpAlaGlyGlyGlySerLeuTrpAsnValLeuHisGlyValGlu 120
DB 463 TGTCTTGTGATGAATATGCTGAAGGGGGCTTTATATATATAGTGTGATGGTGAA 522
QY 121 ProLeuProTyTyrThrAlaAlaHisAlaMetSerTrpCysLeuGlnCysSerGlnGly 140
DB 523 CCATTGCCATATTATACGCTGCCACGACATGATGATGGTGTTCAGTGTTCACAGGA 582
QY 141 ValAlaTrpLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysProPro 160
DB 583 GTGCTTATCTTTCACAGCATCAACCAAGCGTTAATTCACAGGAGCTGAAACCA 642
QY 161 AsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCys 180
DB 643 AACTTACTGCTGTGTGACAGGGGAGACAGTTCT-AAAATTGTGATTTGGGAGACCTGT 701
QY 181 AspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGluVal 200
DB 702 GGACTTCAGACACATGACCAATTAACAAGGGGAGTCTCTGTGAGAGCACCTGGAAT 761
QY 201 PheGluGlySerAsnTrpSerGluLysCysAspValPheSerTrpGlyIleIleLeuTrp 220
DB 762 TTTGGAGGAGTAATTAACGGGAAAAATGACGCTTTCAGCTGGGGAATTTTCTTTGG 821

QY 420 G|YThrGluProGlyGlnValSerSerArgSerSerProSerValArgMetIleThr 439
 DB 541 GGGACGAGAACCTGTCAGAGGACCGGTCATCCACCTTGTCTCAGAAATGATCACT 600
 QY 440 ThrSerGlyProThrSerGlyValProThrArgSerHisProThrProAspSer 459
 DB 601 ACCTCAAGACCACTCAAGAAAGCACTCGCATATCACCCTGAGCCCTGATATTC 660
 QY 460 Thr 460
 DB 661 ACA 663
 RESULT 12 646 bp mRNA linear EST 11-DEC-2001
 LOCUS BJ073883
 DEFINITION BJ073883 NIBB Mochii normalized Xenopus tailbud library Xenopus laevis cDNA clone XLI01n05 5', mRNA sequence.
 ACCESSION BJ073883 GI:17504072
 VERSION BJ073883.1 GI:17504072
 KEYWORDS EST.
 SOURCE African clawed frog.
 ORGANISM Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae; Xenopodinae; Xenopus.
 1 (bases 1 to 646)
 Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-I,T. and Kohara Y.
 REFERENCE Expressed genes in X. laevis embryo
 AUTHORS Unpublished (2001)
 JOURNAL Contact: Tadao Shin-I
 COMMENT Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshini@genes.nig.ac.jp.
 FEATURES
 source Location/Qualifiers
 1..646
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 /clone="XLI01n05"
 /clone_lib="NIBB Mochii normalized Xenopus tailbud library"
 /cissue_type="whole embryo"
 /dev_stage="stage 25"
 /note="vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; cDNAs were oligo-dT primed and directionally cloned. Staging according to Nieuwkoop and Faber. Library is subtracted and was constructed by N. Garrett and A.M. Zorn, (Wellcome/CRC Institute)."
 BASE COUNT 190 a 132 c 143 g 181 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,07e-92 Length: 646
 Score: 1086.00 Matches: 203
 Percent Similarity: 96.26% Conservatives: 3
 Best Local Similarity: 94.86% Mismatches: 8
 Query Match: 36.03% Indels: 1
 DB: 13 Gaps: 0
 US-09-830-144-2 (1-579) x BJ073883 (1-646)
 QY 114 AsnValIeuHisGlyAlaGluProLeuProTyrThrAlaAlaHisAlaMetSerTyr 133
 DB 4 AATGTTTGCATGAGCTGAACCTTGCCTTACAT-ACGTGCTCCATGCAATGAGTGG 62
 QY 134 CysIeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaIleu 153
 DB 63 TGTTTACAAATGTGCCCAAGAGTTGCATATTTACATAGCATGAAGCCAAAGGCTTGATT 122

QY 154 HisArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIle 173
 DB 123 CACAGGAGCTCAACCAACCAACTTGTCTGGTACTGAGGCACTGTTCTTAAGATT 182
 QY 174 CysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAla 193
 DB 183 TGTGACTTTGGTACAGCTGTGATATTACAGCTCAGACTGACTATATAACAAGGAAGTGA 242
 QY 194 AlaTrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPhe 213
 DB 243 GCATGATGTGCTCCAGAAAGTTTGAAGTAGCAACTACAGCCAAAAATGTGACGTGTT 302
 QY 214 SerTrpGlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGly 233
 DB 303 AGTTGGGCAATTATCTTTGGGAAGTAATACCGGAAGAAACCTTTCATGAATTTGGT 352
 QY 234 GlyProAlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLys 253
 DB 363 GGTCCAGGCTTCCTATATATGTGGCTGTTCACAATGTACTCGGCCACATTATTA 422
 QY 254 AsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGln 273
 DB 423 AATTGGCTTAAGCTATTGAAGCTTATGACTGCTGCTGTCGCAAAAGATCCCCACAA 482
 QY 274 ArgProSerMetGluGluIleValLysIleMetThrHisLeuMetArgGlyTrpPheProGly 293
 DB 483 AGACCTTCATGAGGAGGATTTGTAAGATATGACATATTAAGCAAGTATTTCCCTGGA 542
 QY 294 AlaAspGluProLeuGlnTyrProCysGlnTyrSerAspGluGlnSerAsnSerAla 313
 DB 543 GCAGACGTTCTTACAGATATCTTGTCACTGACTGATATGAAGGCAACCAATTCGCA 602
 QY 314 ThrSerThrGlySerPheMetAspIleAlaSerThrAsnThr 327
 DB 603 ACAAGTACAGGTCCTGACTGATATTACATCTACTACT 644
 RESULT 13 604 bp mRNA linear EST 30-JAN-2002
 LOCUS BM426610
 DEFINITION pgf2n.pk004.d15 Normalized Chicken Abdominal Fat Library (pgf2n)
 Gallus gallus cDNA clone pgf2n.pk004.d15 5' similar to
 emb|CB87605.1 (Alu21964) d154614.1.4 (mltogen-activated protein
 kinase kinase kinase 7 (TGF-beta activated kinase id (TRAK1))) [Homo
 sapiens], mRNA sequence.
 ACCESSION BM426610
 VERSION BM426610.1 GI:18431087
 KEYWORDS EST.
 SOURCE chicken.
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianidae; Gallus.
 REFERENCE 1 (bases 1 to 604)
 COGBURN,L.A., MORGAN,R. and BURNSIDE,J.
 ESTs from Normalized Chicken fat cDNA library-USDA/IFARS Animal
 TITLE Genome Project
 JOURNAL Unpublished (2002)
 COMMENT Contact: Larry A. Coghburn
 University of Delaware
 Townsend Hall, Newark, DE 19717, USA
 Tel: 302-831-1335
 Fax: 302-831-2822
 Email: coghburn@udel.edu, www.chickest.udel.edu.
 FEATURES
 source Location/Qualifiers
 1..604
 /organism="Gallus gallus"
 /strain="Commercial broiler, Ottawa Research Centre, leghorn"
 /db_xref="taxon:9031"
 /clone="pgf2n.pk004.d15"
 /clone_lib="Normalized Chicken Abdominal Fat Library (pgf2n)"
 /sex="Male and Female"


```

/tissue_type="Abdominal Fat"
/dev_stage="Embryonic (dl8,dl9); post-hatch (dl.w3,w7,w9,w16,lyr)"
/lab_host="E. coli EMDH10B"
/notes="Vector: pCMVSPORT6; Library made from equivalent pools of total RNA isolated from each developmental age (across strains); Single pass sequencing from 5'-end"
BASE COUNT 167 a 125 c 140 g 167 t 5 others
ORIGIN

```

```

Alignment Scores:      8.52e-92      Length:      604
Pred. No.:            1079.00      Matches:      196
Score:                98.51%      Conservative:  2
Percent Similarity:   97.51%      Mismatches:   3
Best Local Similarity: 35.80%      Indels:       0
Query Match:         13          Gaps:         0
DB:

```

US-09-830-144-2 (1-579) x BM426610 (1-604)

```

QY 105 GluTyrAlaGluGlySerLeuTyrAsnValLeuHisGlyAlaGluProLeuProTyr 124
|||
Db 2 GAGTATGCTCAAGGAGTCTCTGTACAAATGTGCTGATGCTGAACCTCTGCCTCAT 61
|||
QY 125 TyrThrAlaAlaHisAlaMetSerTyrCysLeuGlnCysSerGlnGlyValAlaTyrLeu 144
|||
Db 62 TATACTGCTGCACACCAATGAGTGTGGTGTACAGTGTCCCAAGGAGTGGCATACTTT 121
|||
QY 145 HisSerMetGlnProLysAlaLeuHisArgAspLeuLysProAsnLeuLeuLeu 164
|||
Db 122 CACAGTATGAACCAAGGCCCTAATTCACAGAGACTGAAACCACCAATTTGCTCTTG 181
|||
QY 165 ValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCysAspIleGlnThr 184
|||
Db 182 GTAGTGGGGGACAGTTCTAAAGATCTGTGATTTTGGTACAGCCTGTGATTTCAAACA 241
|||
QY 185 HisMetThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGluValPheGluGlySer 204
|||
Db 242 CACATGACCAACAATAAGGGAAGTGTGCTTGGATGGACCTGAAGTTTTTGAAGGTAGC 301
|||
QY 205 AsnTyrSerGluLysCysAspValPheSerTyrGlyIleLeuTyrGluValIleThr 224
|||
Db 302 AATTACAGTGAANAATGTGAGTTTTCAGTTGGGGTATTATTCTTTGGGAGTGAATCAC 361
|||
QY 225 ArgArgLysProPheAspGluIleGlyProAlaPheArgIleMetTrpAlaValHis 244
|||
Db 362 CGTAGGAACCTTTGATGAGATTGGTGTCCAGCTTTCGCAATAATGTGGCAGTTTAC 421
|||
QY 245 AsnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGluSerLeuMetThr 264
|||
Db 422 AATGGTACTCGACCCACCTGATCAAAACTTACCTAAACCAATTGAGAGTTTAATGACC 481
|||
QY 265 ArgCysTrpSerLysAspProSerGlnArgProSerMetClnGluIleValLysIleMet 284
|||
Db 482 CGCTGTGGTCCAAAGGATCCCTCACACGACCTTCCATGAGGAAATTTGTTAAATAATG 541
|||
QY 285 ThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyrProCysGlnTyr 304
|||
Db 542 ACACACTTGTATGGGTACTTTCCNNNAGCTGATGAACCTCTGCAGTATCTCTGCNAGTAT 601
|||
QY 305 Ser 305
|||
Db 602 TCA 604

```

```

RESULT 14
AL520975 LOCUS AL520975 929 bp mRNA linear EST 13-FEB-2001
DEFINITION AL520975 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB002YM17 5
prime, mRNA sequence.
ACCESSION AL520975
VERSION AL520975.1 GI:12784468
KEYWORDS EST.
SOURCE human.

```

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 929)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
Location/Qualifiers
source 1..929
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DB002YM17"
/clone_lib="LTI_NFL004_NBC2"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 213 a 235 c 249 g 222 t 10 others
ORIGIN

```

```

Alignment Scores:      5.58e-90      Length:      929
Pred. No.:            1063.00      Matches:      204
Score:                100.00%      Conservative:  1
Percent Similarity:   99.51%      Mismatches:   0
Best Local Similarity: 35.27%      Indels:       0
Query Match:         9          Gaps:         0
DB:
US-09-830-144-2 (1-579) x AL520975 (1-929)
QY 1 MetSerThrAlaSerAlaAlaSerSerSerSerAlaGlyGluMetIleGlu 20
Db 305 ATGTCTACAGCCTCTGCGCCTCTCTCTCTCTCTCTCTCGCGGTGATGATCGAA 364
|||
QY 21 AlaProSerGlnValLeuAsnPheGluGluIleAspTyrLysGluIleGluValGluGlu 40
|||
Db 365 GCCCTTCCCAAGTCCCTCACTTTGAAGAAATCGACTACAAGGAGATCGAGGTGGAAGAR 424
|||
QY 41 ValValGlyArgGlyAlaPheGlyValValCysLysAlaLysTrpArgAlaLysAspVal 60
|||
Db 425 GTTGTGGAGAGAGAGCCTTTGGAGTTGTTTGGAAAGCTAAGTGGAGAGACCAAAATGTT 484
|||
QY 61 AlaIleLysGlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
|||
Db 485 GCTATTAACAATAAGAAAGTGAATCTGAGAGAAAGCGTTTATTGTAGAGCTTCGGCAG 544
|||
QY 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnProVal 100
|||
Db 545 TTATCCCGTGTGAACCACTCCTTAATATTGTAAGAGCTTTATGGAGCCTGCTTGAATCCAGTG 604
|||
QY 101 CysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGlyAlaGlu 120
|||
Db 605 TGCTCTGTGATGGAATATGCTGAAGGGGGCTCTTTATATATATGCTGCATGGTGGCTGAA 664
|||
QY 121 ProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeuGlnCysSerGlnGly 140
|||
Db 665 CCATTGCCATATTATCTGCTGCCACCGCAATGAGTTGGTGTTCACAGTGTTCCTCAAGGA 724
|||
QY 141 ValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysProPro 160
|||

```

Db 725 GTGGCTTATCTTACAGCATGCAACCCAAAGCGCTAATTCACAGGAGCCTGAACCAACCA 784
Qy 161 AsnleuleuleuValAlaGlyGlyThrValIleuLysIleCysAspPheGlyThrAlaCys 180
Db 785 AACCTACTGCTGGTTGACGGGGAGACAGTTCTAAATTTGTGATTTTGTGACAGCGCTGT 844
Qy 181 AspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGluVal 200
Db 845 GACATTCAGCACACATGACCAATTAACAAAGGAGGTGCTGTGATGCGMCTGAAGTT 904
Qy 201 PheGluGlySerAsn 205
Db 905 TTGGAAGCTAGTAAT 919
RESULT 15
BG548917 686 bp mRNA linear EST 04-APR-2001
LOCUS 602575750F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4703580 5',
DEFINITION mRNA sequence.
ACCESSION BG548917 GI:13547582
VERSION BG548917.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 686)
AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1542 row: d column: 13
High quality sequence stop: 671.
Location/Qualifiers
1.686
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4703580"
/clone_1ib="NIH_MGC_77"
/lab_host="DH10B (TI phage-resistant)"
/note="Organ: Lung; Vector: pDNR-LIB (Clontech); Site: 1:
SfiI (ggccgctcgcc); Site 2: SfiI (ggccattatggcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGCCATTAATGGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCCGCGCGCGCATG-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.9
kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."
BASE COUNT 202 a 143 c 169 g 171 t 1 others
ORIGIN
Alignment Scores:
Pired. No.: 5,28e-90 Length: 686
Score: 1061.00 Matches: 215
Percent Similarity: 94.78% Conservative: 3
Best Local Similarity: 93.48% Mismatches: 8
Query Match: 35,20% Indels: 5
DB: 12 Gaps: 0
US-09-830-144-2 (1-579) x BG548917 (1-686)
Qy 155 ArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValIleuLysIleCys 174
|||||

Db 4 AGGAACTGAAACCAACCAACTTACTGCTGTAGCTGGGGGACAGTTCTAAAAATTTGT 63
Qy 175 AspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAla 194
Db 64 GATTTGGTACAGCTGTGACATTCAGACACACATGACCAATTAACAAAGGAGAGTGTGCT 123
Qy 195 TrpMetAlaProGluValPheGluGlySerAsnTyrSerGlu-LysCysAspValPhe 214
Db 124 TGGATGGACCTGAAGATTTTGAAGTAGTAATTACAGTAAACCAATGACCTCTTCAG 183
Qy 214 TTPGlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGly 224
Db 184 CTGGGGTATTATTCTTTGGGAAGTATGATACCGCTGGAAACCTTTGATAGATTGGTGG 243
Qy 224 YProAlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProIleLys-A 254
Db 244 CCCAGCTTCCGATCATGTGTGGCTGTTCATTAATGTAATCGNACAGACCTGAATAACAA 303
Qy 254 snLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGln- 273
Db 304 ATTTACCTTAAGCCCATTTGAGAGCGCTGATGACTGTAGTGTCTTAAAGATCTTCCAGG 363
Qy 274 ArgProSerMetGluGluIleValLysIleMetThrHisLeuMetArgTrpPheProGly 293
Db 364 CGCCCTTCAATGAGAGAAATTTGTAAATTAATGACTCACTTGATCGGTACTTTCCAGGA 423
Qy 294 AlaAspGluProLeuGlnTyrProCysGlnTyrSerAspGluGlnSerAsnSerAla 313
Db 424 GCAGATGAGCCATTAACGATATCCTTGCAGTATTCAGATGACAGACAGACCAACTCGCC 483
Qy 314 ThrSerThrGlySerPheMet-AspIleAlaSerThrAsnThrSerAsnLysSerAspTh 333
Db 484 ACCAGTACAGGCTCATTCATGAGCACATTTGCTTCAAAATACAGAGTAACAAAGTGACAC 543
Qy 333 rAsnMetGluGlnValProAlaThrAsnAspThrIleLysArgLeuGlnSerLysLeu 353
Db 544 TAAATATGAGCAAAATTTCTGCGCAAAATGATCTATTAAACCGCTTAGAATCAAC-ATG 602
Qy 353 uLysAsnGlnAlaLysGlnGlnSerGlnSerGlyArgLeuSerLeuGlyValAsnHisG 373
Db 603 GAACAAATCAGGCAACAGCAACAGAGTGAATCTGACGTTTAAGCTTGGAGACCTCCCGTG 662
Qy 373 ySerSerValGluSerLeuPro 380
Db 663 GAGCAGTGTGAGAGCTTGCC 684

Search completed: December 10, 2002, 02:03:51
Job time : 1821.11 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 9, 2002, 23:00:44 ; Search time 61.482 Seconds
(without alignments)
2888.095 Million cell updates/sec

Title: US-09-830-144-2

Perfect score: 3014

Sequence: 1 MSTAASASSSSSSAGEMIE.....QCKKQLEIVRSQQKRGQTS 579

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2_1/USPTO.spool/US09830144/runat_04122002_141354_2283/app.query.fasta_1.1422
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=Blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09830144 @CNG 1.1.44 @runat_04122002_141354_2283 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCUTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3014	100.0	1788	4	US-09-529-279-14 Sequence 14, Appl
2	3014	100.0	2656	2	US-08-685-625A-5 Sequence 5, Appl
3	3014	100.0	2656	4	US-09-529-279-3 Sequence 3, Appl
4	2975	98.7	2443	2	US-08-685-625A-1 Sequence 1, Appl
5	490	16.3	2120	3	US-09-221-235-4 Sequence 4, Appl
6	490	16.3	2120	3	US-09-221-928-4 Sequence 4, Appl
7	490	16.3	2120	3	US-09-221-527-4 Sequence 4, Appl
8	490	16.3	2120	3	US-09-221-236-4 Sequence 4, Appl
9	490	16.3	2120	3	US-09-221-416-4 Sequence 4, Appl
10	490	16.3	2120	4	US-09-221-245-4 Sequence 4, Appl
11	490	16.3	2120	4	US-09-163-115-4 Sequence 4, Appl
12	490	16.3	2120	4	US-09-221-528-4 Sequence 4, Appl

13	490	16.3	2120	4	US-09-593-553-4 Sequence 4, Appl
14	490	16.3	2120	4	US-09-221-237-4 Sequence 4, Appl
15	481	16.0	1365	3	US-09-221-235-6 Sequence 6, Appl
16	481	16.0	1365	3	US-09-221-928-6 Sequence 6, Appl
17	481	16.0	1365	3	US-09-221-527-6 Sequence 6, Appl
18	481	16.0	1365	3	US-09-221-236-6 Sequence 6, Appl
19	481	16.0	1365	3	US-09-221-416-6 Sequence 6, Appl
20	481	16.0	1365	4	US-09-221-245-6 Sequence 6, Appl
21	481	16.0	1365	4	US-09-163-115-6 Sequence 6, Appl
22	481	16.0	1365	4	US-09-221-528-6 Sequence 6, Appl
23	481	16.0	1365	4	US-09-593-553-6 Sequence 6, Appl
24	481	16.0	1365	4	US-09-221-237-6 Sequence 6, Appl
25	418.5	13.9	3025	4	US-09-291-839-1 Sequence 1, Appl
26	418	13.9	2505	4	US-09-291-839-3 Sequence 3, Appl
27	412	13.7	3389	1	US-08-395-580-1 Sequence 1, Appl
28	412	13.7	3426	1	US-08-205-018-1 Sequence 1, Appl
29	398	13.2	2890	1	US-07-928-464-1 Sequence 1, Appl
30	398	13.2	2890	5	PCT-US93-07347-1 Sequence 1, Appl
31	398	13.2	3033	1	US-08-003-311B-1 Sequence 1, Appl
32	398	13.2	3033	1	US-08-261-432-1 Sequence 1, Appl
33	397	13.2	3516	3	US-09-188-930-257 Sequence 257, App
34	384	12.7	1888	3	US-09-188-930-66 Sequence 66, Appl
35	357.5	11.9	2574	4	US-09-142-529-2 Sequence 2, Appl
36	355.5	11.8	3132	3	US-08-167-919A-9 Sequence 9, Appl
37	355.5	11.8	3132	3	US-08-715-106-9 Sequence 9, Appl
38	354	11.7	4508	5	PCT-US93-06251-34 Sequence 34, Appl
39	349.5	11.6	3791	3	US-09-377-310-1 Sequence 1, Appl
40	346	11.5	3133	1	US-08-162-809-1 Sequence 1, Appl
41	345.5	11.5	2442	1	US-08-542-363-3 Sequence 3, Appl
42	345.5	11.5	2442	4	US-09-100-089-3 Sequence 3, Appl
43	345.5	11.5	2442	4	US-09-670-827-3 Sequence 3, Appl
44	345.5	11.5	6226	1	US-08-542-363-1 Sequence 1, Appl
45	345.5	11.5	6226	4	US-09-100-089-1 Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-529-279-14
; Sequence 14, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: TOSHIO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529, 279
; CURRENT FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 14
; LENGTH: 1788
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)..(1776)
US-09-529-279-14

Alignment Scores:
Pred. No.: 8.17e-272 Length: 1788
Score: 3014.00 Matches: 579
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-830-144-2 (1-579) x US-09-529-279-14 (1-1788)

Qy	1	MeerThraIaSerIaIaIaSerSeSerSeSerSeSerIaGIyImetIlelu	20
Db	7	ATGCTACAGCCTCTGCGCCCTCTCTCTCTCTCTCTCTCTCGCGGATGATCGAA	66
Qy	21	AlaProSerGlnValIeuAenPheGlnIuIleAspTyIySgIuIleIuValGlu	40
Db	67	GCCCTTCCGAGGCTCACTTGTGAAGATGACTAACAAGAAATCGAGGTGAAG	126
Qy	41	ValValGlyArgGlyAlaPheGlyValValCysIyAlaIlyTPrArgAlaIysAspVal	60
Db	127	GTTTGTGAAGAGAGCCTTTGGAGTGTGTAAGCTTAAGTGAAGAGCAAAAGATT	186
Qy	61	AlaIleIysGlnIleGluSerGluSerGluValAlaPheIleValGluIeuArgIln	80
Db	187	GCTATTAAACAAATAGAAATGAACTCGAAGAGAAAGCCTTATTGTAGAGTTCCGAG	246
Qy	81	LeuSerArgValaAenHisProAenIleValIyLeuTyGlyAlaCysIeuAenProVal	100
Db	247	TTATCCGCTGGAACCATCTTAATTATGTAAACCTTTATGAGCGCTGTGAATCAATG	306
Qy	101	CysIeuValIuMetGluTyAlaGluGlyGlySerLeuTyAsnValIeuHisGlyAlaGlu	120
Db	307	TGCTTTGATGAGAAATGCTGGAAGGGGCTCTTAAATAGTCTGATGAGTGGTGA	366
Qy	121	ProIeuProTyTyTrThrAlaAlaHisAlaMetSerTyCysIeuGlnCysSerGlnIy	140
Db	367	CCATTGCCATATTACTCTGCTCCCAAGCAATAGTGGTGTTCACAGTTCCTCAAGA	426
Qy	141	ValAlaTyIeuHisSerMetGlnProIyAlaIeuIleHisArgAspLeuIySProPro	160
Db	427	GTGCTTATCTTCACAGACAGCAACCAAGCCCTTATTCACAGGAGCCTGAAACACCA	486
Qy	161	AsnIeuIeuIeuValAlaGlyGlyTyThrValIeuIyAlaCysAspPheGlyThraIaCys	180
Db	487	AACCTACGCTGGTGGAGGGGAGACGTTAAATTTGGATTGGTACACCGT	546
Qy	181	AspIleGlnThrHisMetTrpAsnAsnIySgIySerAlaAlaTrpMetAlaProIuVal	200
Db	547	GACATTCAACACACATGACCAATTAACAAGGGAGTGTGTTGATGAGTGCACCTGAATT	606
Qy	201	PheGluGlySerAsnTyIySerGluIySAspValPheSerTyPglyIleIleuTrp	220
Db	607	TTTGAAGTAGTAAATTACAGTGAAGAAATGACGCTTCAGCTGGGGATATTCTTTGG	666
Qy	221	GluValIleThrArgArgIySProIleAspGluIleGlyGlyProAlaPheArgIleMet	240
Db	667	GAAAGTAAAGCGCTCGAAACCCCTTTGAGAAATTGGGCGCCACTTTCCGAATCATG	726
Qy	241	TrpAlaValHisAsnGlyTyThrArgProProIeuIleIyAsnIeuProIySProIleGlu	260
Db	727	TGGCGCTTCATAATGCTACTCCAGCACCACTATAAAAAATTTACTTAAGCCATTAG	786
Qy	261	SerIeuMetThrArgCysTrpSerIySAspProSerGlnArgProSerMetGluIuIle	280
Db	787	AGCGTGAATGACTCGTGTGTGGTCTTAAAGATCTTCCACGCGCTTCATGAGGAATAAT	846
Qy	281	ValIySleIuMetThrHisIeuMetArgTyIyPheProGlyAlaAspGluProIeuGlnTy	300
Db	847	GTAAGAAATAAAGACTCACTTGATGCGGTACTTTCACAGAGCAATGACCATTAACGAT	906
Qy	301	ProCysGlnTyIySerAspGluGlyGlnSerAsnSerAlaThrSerThrGlySerPheMet	320
Db	907	CTTGTCAGTATTCAGATGAAGAGACAGACACTCGCCACCAAGTACAGGCTCATTCATG	966
Qy	321	AspIleAlaSerThrAsnThrSerAsnIySAspThrAsnMetGluIuValProAla	340
Db	967	GACATTGCTCTTACAAATACGAGTAACAAAGTGAACACTTAATGAGCAACTTCTCC	1026
Qy	341	ThraAsnAspThrIleIySAsyGluGluSerIyIeuIeuIySAsnGlnAlaIySgIuGln	360
Db	1027	ACAAATGATCTATTAAAGCGCTTGAATCAAAATTTGTAATAATACAGCAAGCAACAG	1086

```

OY 361 SerLysSerGlyThrLeuSerLeuGlyValAspThrLysSerSerValGlySerLeuPro 380
Db 1087 AGTGAATCTTGAGACGTTTAAGCTTTGGAGCCTCCCATGGAGACGTCGTGAGACGCTGCC 1146
OY 381 ProThrSerGlyGlyLysArgMetSerAlaAspMetSerGluLeuAlaArgLLea 400
Db 1147 CCAACCTCTGAGGGCAAGAGATGATGCTGCATGTCATGTCGAATATGAGAGCTAGGATCCGC 1206
OY 401 AlaThrThrGlyValGlnPheArgAlaGlySerLLeuAlaSerLLeuValThrGly 420
Db 1207 GCAACCAACAGGACCAAGGACGACCAACGATGATTCATCCAAAGACTTGAACGTGAACGGA 1266
OY 421 ThrGluProGlyGlnValSerSerArgSerSerProSerValArgMetLLeuThrThr 440
Db 1267 ACAAGACCTGTCAGGTGAGACGATAGGTATCATGATCCCAAGTGTGAGATGATTTACTACC 1328
OY 441 SerGlyProThrSerGlyLysProThrArgSerHisProTyrThrProAspAspSerThr 460
Db 1327 TCAGAGCAACCACTCAGAAAAGCCCACTCGAAGTCATCCATGAGACCCCTGATGATTCGACA 1386
OY 461 AspThrAsnGlySerAspAsnSerLLeuPheMetAlaTyrLeuThrLeuAspHisGlnLeu 480
Db 1387 GATACCAATGGATGATAGATTAATCATCCATCCCAATGGCTTATCTTACACTGATCACCACATA 1446
OY 481 GlnProLeuAlaProCysProAsnSerLysGlySerMetAlaValPheGlnGlnHisCys 500
Db 1447 CAGCCTCTAGACCGGTGCCCAAACTCCAAAGATCTATGGCAGTGTGGAAGACGATGTT 1506
OY 501 LysMetAlaGlnGlyThrMetLysValGlnThrGluLeuAlaLeuLeuLeuGlnArgLys 520
Db 1507 AAAATGGCACAAGATATATGAAAGTTCAAACAGAAATTCGATTTGTTTACAGAGAAAG 1566
OY 521 GlnGluLeuValAlaGluLeuAspGlnAspGlyLysAspGlnGlnAsnThrSerArgLeu 540
Db 1567 CAAGACTAGTGTGCAGAACCTGCAGACCGAGATGAAAGACACAGCAAAATATCATCTGCCTG 1626
OY 541 ValGlnGlnHisLysLysLeuLeuAspGlnAsnLysSerLeuSerThrTyrTyrGlnGln 560
Db 1627 GTAACGGAACCTAAAGAGCTTTATGATGAAACAAAGGCTTTCTTACTTACTACACAGCAA 1686
OY 561 CysLysLysGlnLeuGlnValLLeuArgSerGlnGlnGlnLysArgGlnGlyThrSer 579
Db 1687 TGCAAAACCAACTGACGATGATCATCAGAAATCAGACGACCAAAACGACAGGACCTTCA 1743

RESULT 2
US-08-685-625A-5
; Sequence 5, Application US/08685625A
; Patent No. 5945301
; GENERAL INFORMATION:
; APPLICANT: UENO, Naoto
; APPLICANT: MATSUMOTO, Kunihito
; APPLICANT: IRIE, Kenji
; TITLE OF INVENTION: NOVEL KINASE IN TGF-BETA FAMILY SIGNAL
; TITLE OF INVENTION: TRANSDUCTION SYSTEM
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATTHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/685,625A
; FILING DATE: 24-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-253549

```



```
Db 1803 GTACAGGAACATAAAAGCTTTAGATGAAACAAAAAGCCCTTTCTACTTACTACAGCAA 1862
Qy 561 CysLysLysGlnLeuGluValIleArgSerGlnGlnLysArgGlnGlyThrSer 579
Db 1863 TGCAAAAAACACTAGAGGTTCATCAGAGTCATCAGAGTCAGAGAGTACAGAGAACACACAGGCACTTCA 1919

RESULT 4
US-08-685-625A-1
; Sequence 1, Application US/08685625A
; Patent No. 5945301
; GENERAL INFORMATION:
; APPLICANT: UENO, Naoto
; APPLICANT: MATSUMOTO, Kunihiko
; APPLICANT: IRIE, Kenji
; TITLE OF INVENTION: NOVEL KINASE IN TGF-BETA FAMILY SIGNAL
; TITLE OF INVENTION: TRANSDUCTION SYSTEM
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22131-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/685,625A
; FILING DATE: 24-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-253549
; FILING DATE: 29-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Meuth, Donna M.
; REGISTRATION NUMBER: 36,607
; REFERENCE/DOCKET NUMBER: 001560-267
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2443 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 157..1893
US-08-685-625A-1

Alignment Scores:
Pred. No.: 5,926-268 Length: 2443
Score: 2975.00 Matches: 572
Percent Similarity: 98.79% Conservative: 0
Best Local Similarity: 98.79% Mismatches: 7
Query Match: 98.71% Indels: 0
DB: 2 Gaps: 0

US-09-830-144-2 (1-579) x US-08-685-625A-1 (1-2443)

Qy 1 MetSerThrAlaSerAlaAlaSerSerSerSerSerSerSerSerAlaGlyUleGlu 20
Db 157 ATGTGACAGCCCTCGCGCCCTCGTCCCTCCCTCGTCTTCTGCGCAGTCAGATCGAA 216
Qy 21 AlaProSerGlnValLeuAsnPheGluGluIleAspTyrLysGluIleGluValGluGlu 40
Db 217 GCGCGCTGCGAGTCTCTGAACCTTCGAGAGATCGACTACAGAGGATCGAGGTGGAAGAG 276
```

QY 401 AlaThrThrGlyAsnGlyGlnProArgArgSerIleGlnAspLeuThrValThrGly 420
 DB 1357 GCGACTGCAAGGTAACGGCAACCAAGGGGTATCATCCAAAGCTTACGTACTGGG 1416
 QY 421 ThrGluProGlyGlnValSerSerArgSerSerProSerValArgMetIleThrThr 440
 DB 1417 ACAGAACTGGTGCAGGTGAGCGCGTATCATCCAGCCCTAGTGCAGATGATCACTACC 1476
 QY 441 SerGlyProThrSerGlyLeuSerProThrArgSerHisProThrThrProAspSerThr 460
 DB 1477 TCAGGACCAACTCAGAGAGCCAGCTGCGATCCATCCATGATCCATGATTCACCA 1536
 QY 461 AspThrAsnGlySerAspAsnSerIleProMetAlaTyrLeuThrLeuAspHisGlnLeu 480
 DB 1537 GACACCAATGGCTCAGATTAATCTCATCCATGGCGTATCTACACTGGATCAGAGCTA 1596
 QY 481 GlnProLeuAlaProCysProAsnSerIleGlySerMetAlaValPheGlnGlnHisCys 500
 DB 1597 CAGCCCTGAGCGCGCGCCCAACTCCAAAGATCCATGGCGTGTTCGAAACAGACTGT 1656
 QY 501 LysMetAlaGlnGluTyrMetLysValGlnThrGlnIleAlaLeuLeuGlnArgLys 520
 DB 1657 AAATGACACAGAGATATTTGAAGTTCAAACCGAAATCGCATTTGTACTACAGAGAAAG 1716
 QY 521 GlnGluLeuValAlaGlnLeuAspGlnAspGluLysAspGlnGlnAsnThrSerArgLeu 540
 DB 1717 CAAGAACTGTTGCAAGATTGGACCGAGATGAAAGACCAAGCAAAATATCATCTCGTGTG 1776
 QY 541 ValGlnGlnHisLysLysLeuLeuAspGluLysSerLeuSerThrTyrTyrGlnGln 560
 DB 1777 GTTACAGGAACATPAAAAAGCTTTTATGTAAGAAACAAAAGCCCTTTTATTAATTAACAGCAA 1836
 QY 561 CysLysLysGlnLeuGlnValIleArgSerGlnGlnGlnLysArgGlnGlyThrSer 579
 DB 1837 TGCAGAAACCACTAAGGTCAATCAGAGCCAAACAGAGAAAGCAAGAGCACTTCA 1893

RESULT 5
 US-09-221-235-4
 ; Sequence 4, Application US/09221235
 ; Patent No. 6043040
 ; GENERAL INFORMATION:
 ; APPLICANT: Acton, Susan
 ; TITLE OF INVENTION: NOVEL CSAK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
 ; FILE REFERENCE: MNI-050
 ; CURRENT APPLICATION NUMBER: US/09/221,235
 ; CURRENT FILING DATE: 1998-12-28
 ; EARLIER APPLICATION NUMBER: 09/163,115
 ; EARLIER FILING DATE:
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 4
 ; LENGTH: 2120
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (47)..(1411)
 ; US-09-221-235-4

Alignment Scores:
 Pred. No.: 4,97e-36 Length: 2120
 Score: 490.00 Matches: 138
 Percent Similarity: 48.79% Conservative: 84
 Best Local Similarity: 30.33% Mismatches: 159
 Query Match: 16.26% Indels: 74
 DB: 3 Gaps: 19

US-09-830-144-2 (1-579) x US-09-221-235-4 (1-2120)

QY 10 SerSerSerSerAlaGlyGluMetIle-----GlnAlaProSerGlnVal 25
 DB 2 TCGACCAACGCGTCCGGTGAAGTAAATCTTTGTCTATATGAGATGTGCTCTCGGT 61

QY 26 LeuAsnPheGlnGluIleAspTyrLysGlnIleGluValGluGluValAlaGlyArgGly 45
 DB 62 GCGCTCTTGTGCAATTAATTTGATGACTGCACTTTTGTGAAAACCTGCGGTGAGGA 121
 QY 46 AlaPheGlyValAlaCysLysAlaLysTrp-----ArgAlaLysAspValAlaLys 63
 DB 122 AGTTTGGAGTGTATTATGAGCCAAATGATATTCACAGACAGAGAGGTGCTGTAAAG 181
 QY 64 -----GlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
 DB 182 AAGCTCTCAAAATPAGAAAGAGCAGAA-----ATA 214
 QY 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnPro---- 99
 DB 215 CTGAGTCTCTCAGTCACAGAAACATCATCAAGTTTATGAGTAATTTTGAACCTCCC 274
 QY 100 ---ValCysLeuValMetGluTyrAlaGluGluGlySerLeuTyrAsnValLeuHisGly 118
 DB 275 AACTATGGCATTTGTACAGAAATATGCTTCTCTGGATCACTCATATGATTAACATTAACAGT 334
 QY 119 -----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeuGln 136
 DB 335 AACAGAGTGAGAGATG-----GATATGATCACAATTATGACTCTGGCCACTGAT 385
 QY 137 CysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAsp 156
 DB 386 GTACCAAGAGAAAGCATTTATTTACATATGAGAGCTCTCTCAAGGTATTCAAGAGAC 445
 QY 157 LeuLysProProAsnLeuLeuLeuValAlaGlyLysThrValLeuLysIleCysAspPhe 176
 DB 446 CTCAAGTCAGAAAGATGTTTATAGCTGCGATGAG---GTACTGAAGATCTGTGACTTT 502
 QY 177 GlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys-----GlySerAlaAla 194
 DB 503 GGT---GCTCTCGGTCCATTAACATPACACACATGTCCTTGTGTGAACCTTCCCA 559
 QY 195 TrpMetAlaProGluValAlaPheGlnGlySerAsnTyrSerGluLysCysAspValPheSer 214
 DB 560 TGGATGGCTCCAGAAAGTTATCCAGATCTCCCTGTGTCAAAAAGTTGTGACATATTC 619
 QY 215 TrpGlyIleIleLeuTyrGluValIleThrArgArgLysProPheAspGluIleGly 234
 DB 620 TATGTTGTGTTCTCTGGAGATGCTTAACAAGGAGGTCCCTTTAAAGTTTGAAGA 679
 QY 235 ProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgProProLeuIleLys 253
 DB 680 -----TTACAAAGTAGCTTGCTGTGTAGTGAATAAAGAGATTAAACCATTTCCAGC 733
 QY 254 AsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGln 273
 DB 734 AGTTGCCCCAGAAAGTTTGTCTGAACCTTTACATCAGTGTGGAGAGCTGATGCCAAGAAA 793
 QY 274 ArgProSerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGly 293
 DB 794 CGGCATCATTCAGCAAAATCATTTCAATCCTG----- 826
 QY 294 AlaAspGluProLeuGlnTyrProCysGlnTyrSerAspGluGlnSerAsnSerAla 313
 DB 827 -----GAGTCCATGCAAAATGACAGC 847
 QY 314 ThrSerThrLysSerPheMetAspIleAlaSerThrAsnThrSerAsnLysSerAspThr 333
 DB 848 -----AGCTTCTCGACAAAGTGAATCATTTCTACACAAACAAAGCGGAGTGG 895
 QY 334 AsnMetGluGlnValProAlaThrAsnAspThrIleLysArgLeuGluSerLysLeuLeu 353
 DB 896 AGGTGCGAA---ATTGAAGCAACTCTTGAGAGCTTAAGAAACATGAGCGTATCTCAGC 952
 QY 354 LysAsnGlnAlaLysGlnGlnSerGluSerGlyArgLeuSerLeu-----GlyAla 370
 DB 953 TTTAAGGACGAGACTTAAAGAAAGAGAAAGAGTTTAAAGATGTGAGAGCAAAAGCTG 1012
 QY 371 SerHisGlySerSerValGluSerLeuProProThrSer----- 383


```

Db 1013 ACAGAGAGTCCACACCCCGCTTCTCTGCTCTGCTCAAGAATGTCTGAGGAGTCT 1072
Qy 384 -----GluGlyLysArgMetSerAlaAspMetSerGluLeuGluAlaAarglleAlaAa 401
Db 1073 TACTTTGAATCTAAACAGAGGAGTCAACAGTGCAGAGATGTCATGTTCAGATCACAGCA 1132
Qy 402 ThrThr---GlyAenGlyGlnProArgArgSerIleGlnAspLeuThrValThrGly 420
Db 1133 ACAAGTAAACGGGAGGCGCATGGCATGAACCAAGTCTGCAGGCCATGCTGTGATGGGC 1192
Qy 421 ThrGluProGlyGlnValSerArgSerSerSerProSerVal 435
Db 1193 TTT-----GGGGATATCTTCTCAATGAACAAAGCAGGAGCTGTG 1231

```

RESULT 6

```

US-09-221-928-4
; Sequence 4, Application US/09221928
; Patent No. 6121030
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,928
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2120
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (47)..(1411)
US-09-221-928-4

```

Alignment Scores:

```

Pred. No.: 4,97e-36 Length: 2120
Score: 490.00 Matches: 138
Percent Similarity: 48.79% Conservative: 84
Best Local Similarity: 30.33% Mismatches: 159
Query Match: 16.28% Indels: 74
DB: 3 Gaps: 19

US-09-830-144-2 (1-579) x US-09-221-928-4 (1-2120)

Qy 10 SerSerSerSerSerAlaGlyGluMetIle-----GluAlaProSerGlnVal 25
Db 2 TCGACCCAGCGTCCGGTGGGAAGTATAATCTTTGTCTATTATGATGTCGTCTCTCGT 61
Qy 26 LeuAsnPheGluGluIleAspTyrLysGluIleGluValGluGluValValGlyArgGly 45
Db 62 GCCTCTTTGTGCAATTAATTTGATGACTTGCAGTTTTTTTGAACCTGCGGTGGAGA 121
Qy 46 AlaPheGlyValValCysLysAlaLysTrp-----ArgAlaLysAspValAlaIleLys 63
Db 122 AGTTTGGAGTGTATTATCGACCAAAATGGATATCACAGACACAGAGGTGGCTGTAAAG 181
Qy 64 -----GlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
Db 182 AAGCTCTCTCAAAATAGAGAAGAGCGGAGAA-----ATA 214
Qy 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnPro----- 99
Db 215 CTCAGTGCTCTCAGTCACAGAAACATCATCCAGTTTTTATGGAGTAATCTTGAACCTCCC 274
Qy 100 ---ValCysLeuValMetGlyTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGly 118
Db 275 AACTATGGCATTGTTCACAGAAATATGCTTCTCTCGGATCACCTATGATTACATTAACAGT 334
Qy 119 -----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeuGln 136

```

```

Db 335 AACAAAGTGAGGAGATG-----CATATGGATCACATTATGACCTGGGCCACTGAT 385
Qy 137 CysSerGlnGlyValAlaLysLeuHisSerMetGlnProLysAlaLeuIleHisArgAsp 156
Db 386 GTAGCCAAAGGAATGCAATTAATTTACATATGAGGCTCTGTCAAGGTGATTCACAGAGAC 445
Qy 157 LeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPhe 176
Db 446 CTCAAAGTCAAGAAAGTGTGTATAGCTGCTGATGGA---GTACTGAAGATCTGTGACTTT 502
Qy 177 GlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys-----GlySerAlaAla 194
Db 503 GGT---GCCTCTCGGTCCATAACCATACACACATGCTCTTGGTTGGAACCTTTCCCA 559
Qy 195 TrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSer 214
Db 560 TGGATGGCTCCAGAGTTATCCAGAGTCTCCCTGTGTGAGAACTTGTGACACATATTCC 619
Qy 215 TrpGlyIleIleLeuTyrGluValIleThrArgArgLysProPheAspGluIleGlyGly 234
Db 620 TATGTGTGGTTCCTGGAGATGCTAACAAAGGAGGTCCCTTTTAAAGGTTTGGAAGGA 679
Qy 235 ProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgProProLeuIleLys 253
Db 680 -----TTACAAGTAGCTTGGCTTGTAGTGGAAAAAACGAGAGATTAAACATTCCAAAGC 733
Qy 254 AsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGln 273
Db 734 AGTTGCCCCAGAGTTTTTGTGTGAACCTGTATCATCAGTGTGGGAGCTGATGCCAAGAAA 793
Qy 274 ArgProSerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGly 293
Db 794 CGGCATCATTTCAAGCAAAATCATTTCAATCCTG----- 826
Qy 294 AlaAspGluProLeuGlnTyrProCysGlnTyrSerAspGluGlyGlnSerAsnSerAla 313
Db 827 -----GAGTCCATGTCAAATGACACG 847
Qy 314 ThrSerThrGlySerPheMetAspIleAlaSerThrAsnThrSerAsnLysSerAspThr 333
Db 848 -----AGCCTTCTCTGCAAGTGTAACTCATTTCTACACAACAAGCGGAGTGG 895
Qy 334 AsnMetGluGlnValProAlaThrAsnAspThrIleLysArgLeuGluSerLysLeuLeu 353
Db 896 AGTGCAGAA---ATTGAGGCAACTCTTTGAGGCGCTAAAGAAACTAGAGCGTGTCTCAGC 952
Qy 354 LysAsnGlnAlaLysGlnSerGluSerGlyArgLeuSerLeu-----GlyAla 370
Db 953 TTTTAAGGAGCAGGAGCTTTAAAGAACGAGAAAGACGTTTAAAGATGTGGGAGCAAAAGCTG 1012
Qy 371 SerHisGlySerSerValGluSerLeuProProThrSer----- 383
Db 1013 ACAGAGCAGTCCCAACACCCCGCTTCTTTCCTCTTGTGCTGCAAGAATGTCTGAGGAGTCT 1072
Qy 384 -----GluGlyLysArgMetSerAlaAspMetSerGluIleGluAlaAarglleAlaAa 401
Db 1073 TACTTTGAATCTAAACAGAGGAGTCAACAGTGCAGAGATGTCATGTTCAGATCACAGCA 1132
Qy 402 ThrThr---GlyAsnGlyGlnProArgArgSerIleGlnAspLeuThrValThrGly 420
Db 1133 ACAAGTAAACGGGAGGCGCATGGCATGAACCCCAAGTCTGCAGGCCATGATGTGTGATGGGC 1192
Qy 421 ThrGluProGlyGlnValSerArgSerSerSerProSerVal 435
Db 1193 TTT-----GGGGATATCTTCTCAATGAACAAAGCAGGAGCTGTG 1231

```

RESULT 7

```

US-09-221-527-4
; Sequence 4, Application US/09221527
; Patent No. 6146832
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan

```


Best Local Similarity:	30.33%	Mismatches:	159
Query Match:	16.26%	Indels:	74
DB:	3	Gaps:	19

US-09-830-144-2 (1-579) x US-09-221-236-4 (1-2120)

Qy	10	SerSerSerSerSerAlaGlyGluMeIle-----GluAlaProSerGlnVal	25
Db	2	CGACCCACCGGTCGCGTGAAGTATAATACTTTGTCATTATGAGATGCTCTCTCGGT	61
Qy	26	LeuAsnPheGluIleAspTyrLysGluIleGluValGluGluValGlyArgGly	45
Db	62	GCCTCCCTTGTGCAATTAATTTGATGACTTCGAGTTTTTGAAAATCGCGGTGAGGA	121

QY 46 AlaPheGlyValValCysLysAlaLysTrp-----ArgAlaLysAspValAlaIleLys 63
 :::||||| ||| :::||||| ||| :::||||| ||| :
Db 122 AGTTTGGAGTGTTTATCGAGCCAAATTGGATTCACAGGCAGGAGGTGGCTGTTAAG 181

Qy 64 -----GlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
 :::||||| |||:::|
Db 182 AAGTCCTCAAAATAGAGAAAGAGGCAGAA-----ATA 214

Qy 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnPro--- 99

D8 215 CTCAAGTGCTCCTCAGTCACAGAAACATCATCCAGTTTATGGAGTAATTCTTTGAACCTCCC
QY 100 ---ValCysIeuValMetGluTyrAlaGluGlySerLeuTyrAsnValLeuHisGly

D6	275	AACTATGGCATTGTCACAGATATGCTTCTCTGGGATCACTCTATGATTACATTAAACAGT	334
		
Qy	119	-----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrrCysLeuGln	136

Db 335 AACAGAGTGGAGATG-----GATATGGATCACATTATGACCTGGGCCACTGAT 385

Qv 137 CvesSerg[nc]lWla[Ala]tYrLeuH[SerMetC]nBr[LeuMetC]nLeuT[Ala]c[Ala]c[Ala]c 156

47 386 GTAGCCAAGGAATGCATTATTTACATATGGAGGCTCTGTCAAGGTGATTCCACAGAGAC 445

Qy
157 LeuLysProProAsnLeuLeuValAlaGLYThrValLeuLysIleCysAppPhe 176
||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db
446 CTCAGTCAAGAACGTGTTATAGCTGCTGATGGA---GTACTGAAGATCTGTGACTTT 502

Qy 177 GlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys-----GlySerAlaAla 194
||| ||| ||| ||| ||| ||| ||| |||
Db 503 GGT---GCCTCTCGGTTCCATAACCATCAACAACACATGTCCTGGTGGAACTTTCCCCA 559

Qy 195 TrpMetAlaproGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSer 214
Db TGAATCGGGCCACAACTTTATCATCAAGTCGCCGCTGTCTCTGACTCTGCACGATTTCG 218

Oy 215 TrpGlyIleileLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGly 2344
 :::||| ::::::: ||||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
220 1GGATGGCTCCAGAAAGTATTCCAGAGAGTCTCCCTGTGTGCAGATAACITGGTCACATAATCC 819

Db 620 TATGGTGTGGTCTCTGGGAGATGCTAACAGGAGAGGTCCCTTTTAAAGGTTTGGAGGA 679
 Qy 235 ProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgProProLeuIleLys 253

Db 680 -----TTACAAGTAGCTTGGCTTGTACTGGAAAAAACGAGAGATTAAACCATTCCAAGC 733
Qy 254 AsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGln 273

Db 734 AGTTGCCCCAGAGTTTGTGAACTGTATACATCAGTGTGGGAAGCTGATGCCAAGAAA 793

CY 274 AGTGTGCTGGGCGGAGGATTTTCATCTG-----826
||||| : : : : : ||| : :
||| : : : : : ||| : :
Db 794 CGGCATCATTTCAAGCAAATCATTCAATCTG-----826

Qy 294 AlaAspGluProLeuGlnTyrProCysGlnTyrSerAspGluGlyGlnSerAsnSerAla 313
|||
Db 827 -----GAGTCCATGTCAAAATGACACG 847

Qy 314 ThrSerThrGlySerPheMetAspIleAlaSerThrAsnThrSerAsnLysSerAspThr 333
 ||| ||| : : : : : ||| ||| : : : : :
 Db 848 -----AGCCTTCTGTACAAGTGTAATTCATTCTACACAACAAGCGGGACTGG 895

Qy	334	AsnMetGluGlnValProAlaThrAsnAspThrIleLysArgLeuGluSerLysLeuLeu	353
Db	896	AGTGCCAA--ATTGAGGCAACTCTTGAGAGCGCTAAAGAAATAGAGCGTATCTACGC	952
Qy	354	LysAsnGlnAlaLysGlnSerGluSerGlyArgLeuSerLeu-----GlyAla	370
Db	953	TTTAAGGAGCAGGAGCTTTAAAGAACGAGAAAGACGTTTAAAGATGTGGGAGCAAAAGCTG	1012
Qy	371	SerHisGlySerSerValGluSerLeuProThrSer-----	383
Db	1013	ACAGAGCAGTCCAAACACCCCGCTTCTTGCCCTTTGCTGCAAGAATGTCGAGGAGTCT	1072
Qy	384	-----GluGlyLysArgMetSerAlaAspMetSerGluIleGluAlaArgIleAlaAla	401
Db	1073	TACTTTGAATCTAAACACAGAGGAGTCAAAACAGTGCAGAGAGATGTCATGTCAGATCACACA	1133
Qy	402	ThrThr---GlyAsnGlyGlnProArgArgArgSerIleGluAspLeuThrValThrGly	420
Db	1133	ACAAGTAACGGGAGGGCCATGCGATGAACCCAAAGTCTCGAGCCCATGATGCTGATGGC	1193
Qy	421	ThrGluProGlyGlnValSerSerArgSerSerProSerVal	435
Db	1193	TTT-----GGGATATCTTCTCAATGAACAAAGCAGGAGCGTGTG	1231

RESULT 9

```

US-09-221-416-4
; Sequence 4, Application US/09221416
; Patent No. 6153417
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: WNI-050
; CURRENT APPLICATION NUMBER: US/09/221,416
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2120
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (47) ..(1411)
US-09-221-416-4

```

Alignment Scores:

Alignment Scores:		
Pred. No.:	4,97e-36	Length: 2120
Score:	490.00	Matches: 138
Percent Similarity:	48.7%	Conservative: 84
Best Local Similarity:	30.3%	Mismatches: 159
Query Match:	16.26%	Indels: 74
DB:	3	Gaps: 19

US-09-830-144-2 (1-579) X US-09-221-416-4 (1-2120)

Qy	10	SerSerSerSerSerAlaGlyGluMetIle-----GluAlaProSerGlnVal	25
		: : : : : : :	
Db	2	TCGACACACGCGTCGCGTGAAGTAATAACTTTGTCATTATGAGATGTCGTCCTCGGT	61
		: : : : : : :	
Qy	26	LeuAsnPheGluIleAspTyrLysGluIleGluValGluGluValValGlyArgGly	45
		: : : : : : :	
Db	62	GCCTCCTTTGTGCAAATTAATTTGATGACTTCGAGTTTTTGAACACTGCGGTGGAGGA	121
		: : : : : : :	
Qy	46	AlaPheGlyValValCysLysAlaLysTrp-----ArgAlaLysAspValAlaIleLys	63
		: : : : : : :	
Db	122	AGTTTGGGAGTGTTTATCGAGCCAAATGATATCAGCAGCAAGGAGGTGGCTGTAAAG	181
		: : : : : : :	
Qy	64	-----GlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln	80
		: : : : : : :	
Db	182	AAGTCCTCTCAAAATAGAGAAGAGGCGGAGAA-----ATA	214


```
QY 177 GlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys-----GlySerAlaAla 194
Db 503 GGT---GCCTCTCGGTTCCTAATACACACACACATGCTCTGGTGGAACTTTCCCA 559

QY 195 TrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSer 214
Db 560 TGGATGCTCCAGAAAGTTATCCAGAGTCTCCCTGTGTCAGAAACTTGTGACACATATCC 619

QY 215 TrpGlyIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGly 234
Db 620 TATGCTGTGGTCTCTCGGAGATGCTAACAGGGAGGCTCCCTTTAAAGTTTGGAGGA 679

QY 235 ProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgProProLeuIleLys 253
Db 680 -----TTACAAGTAGCTTGGCTTGTAGTGAATAAAGAGAGAGATTAACCATTCACAGC 733

QY 254 AsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGln 273
Db 734 AGTTGCCCCAGAGTTTGTGTAACCTGTTACATCAGTGTTGGGAAGCTGATGCCAAGAA 793

QY 274 ArgProSerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGly 293
Db 794 CGGCCATCATTCAGCAAAATCATTTCAATCTG----- 826

QY 294 AlaAspGluProLeuGlnTyrProCysGlnTyrSerAspGluGlyGlnSerAsnSerAla 313
Db 827 -----GAGTCCATGTCAAAATGACACG 847

QY 314 ThrSerThrGlySerPheMetAspIleAlaSerThrAsnThrSerAsnLysSerAspThr 333
Db 848 -----AGCCTCTCGTCAAGTGAATCACTCACTCTCACAAACAGGGAGTGG 895

QY 334 AsnMetGluGlnValProAlaThrAsnAspThrIleLysArgLeuGluSerLysLeu 353
Db 896 AGGTGCGAA---ATTGAGGCAACTTTCAGAGGCTAAAGAACTAGACGCTGATCTCAGC 952

QY 354 LysAsnGlnAlaLysGlnGlnSerGluSerGlyArgLeuSerLeu-----GlyAla 370
Db 953 TTTAAGGAGCAGAGCTTAAAGAACAGAGAAAGACGTTTAAAGATGTGGGCAAAAGCTG 1012

QY 371 SerHisGlySerValGluSerLeuProProThrSer----- 383
Db 1013 ACAGAGCAGTCCACACCCCGCTTCTTTCCTCTGCTGCTGCAAGAAATGCTGAGGAGTCT 1072

QY 384 -----GluGlyLysArgMetSerAlaAspMetSerGluIleGluAlaAlaAla 401
Db 1073 TACTTTGAATCTAAACAGAGGAGTCAACAGTGCAGAGATGTCATGTCAGATCACAGCA 1132

QY 402 ThrThr---GlyAsnGlyGlnProArgArgSerIleGlnAspLeuThrValThrGly 420
Db 1133 ACAAGTAACGGGAGGCGCCATGGCATGAACCCCAAGTCTCGGCCCATGATGCTGATGGGC 1192

QY 421 ThrGluProGlyGlnValSerSerArgSerSerSerProSerVal 435
Db 1193 TTT-----GGGATATCTTCTCAATGAACAAAGCAGGAGCTGTG 1231

RESULT 11
US-09-163-115-4
; Sequence 4, Application US/09163115A
; Patent No. 6183962
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/163,115A
; CURRENT FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2120
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

```
; NAME/KEY: CDS
; LOCATION: (47)..(1411)
US-09-163-115-4

Alignment Scores:
Pred. No.: 4,97e-36 Length: 2120
Score: 490.00 Matches: 138
Percent Similarity: 48.79% Conservative: 84
Best Local Similarity: 30.33% Mismatches: 159
Query Match: 16.26% Indels: 74
DB: Gaps: 19

US-09-830-144-2 (1-579) x US-09-163-115-4 (1-2120)

QY 10 SerSerSerSerSerAlaGlyGluMetIle-----GluAlaProSerGlnVal 25
Db 2 TCGACCCACCGGTCGGTGGAGATATAATCTTTGTCAATTATGAGATGTCCTCTCGGT 61

QY 26 LeuAsnPheGluGluIleAspTyrLysGluIleValGluValValGlyArgGly 45
Db 62 GCCTCTTTGTGCAAAATTAATTTGATGACTTGCAGTTTGTGAAACTGCGGTGGAGGA 121

QY 46 AlaPheGlyValValCysLysAlaLysTrp-----ArgAlaLysAspValAlaIleLys 63
Db 122 AGTTTGGGAGTGTATTATCGAGCCAAATGGATATCACAGGACAAGGAGGTGGCTGTAAG 181

QY 64 -----GlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
Db 182 AAGCTCTCAAAATAGAGAAAGAGGCAGAA-----ATA 214

QY 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnPro--- 99
Db 215 CTCAGTGTCTCAGTCACAGAAACATCATCCAGTTTATGGAGTAATCTTTGAACCTCCC 274

QY 100 ---ValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGly 118
Db 275 AACTATGCGCATGTGCAGAAATATGCTTCTCTGGCATCACTCTATGATTACATTAAACAGT 334

QY 119 -----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeuGln 136
Db 335 AACAGAGTGTAGGAGATG-----GATATGGATCACATTATGACCTGGGCCACTGAT 385

QY 137 CysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAsp 156
Db 386 GTAGCCAAAGGAAGTATTTATACATATGGAGCTCTCTCAAGGTGATTCACAGAC 445

QY 157 LeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPhe 176
Db 446 CTCAGTCAAGAAACGTTGTTATAGCTGCTGATGGA---GTACTGAAGATCTGTGACTTT 502

QY 177 GlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys-----GlySerAlaAla 194
Db 503 GGT---GCCTCTCGGTTCCTAATACACACACATGCTCTGGTGGAACTTTCCCA 559

QY 195 TrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSer 214
Db 560 TGGATGCTCCAGAAAGTTATCCAGAGTCTCCCTGTGTCAGAAACTTGTGACACATATTC 619

QY 215 TrpGlyIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGly 234
Db 620 TATGCTGTGGTCTCTCGGAGATGCTAACAGGGAGGTCCTTTAAAGTTTGGAGGA 679

QY 235 ProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgProProLeuIleLys 253
Db 680 -----TTACAAGTAGCTTGGCTTGTAGTGAATAAAGAGAGATTAACCATTCACAGC 733

QY 254 AsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGln 273
Db 734 AGTTGCCCCAGAGTTTGTGTAACCTGTTACATCAGTGTTGGGAAGCTGATGCCAAGAA 793

QY 274 ArgProSerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGly 293
Db 794 CGGCCATCATTCAGCAAAATCATTTCAATCTG----- 826
```

QY 294 AlaaspGluProLeuGlnTyrProCysGlnTyrSerAspGluGlyGlnSerAspSerAla 313
DB 827 -----GAGTCATGTCAAATACACG 847
QY 314 ThrSerThrGlySerPheMetAspIleAlaSerThrAsnThrSerAsnLysSerAspThr 333
DB 848 -----AGCCTTCCTGCACAAAGTGAATCACTTCCTACACAAAGCGGAGTGG 895
QY 334 AsnMetGluGlnValProAlaThrAsnAspThrIleLysArgLeuGlnSerLysLeuLeu 353
DB 896 AGGTGGCAA---ATTGAGGCAACTCTTGAGAGGCTAAAGAAACTAGAGCGTGAATTCAGC 952
QY 354 LysAsnGlnAlaLysGlnGlnSerGlnSerGlnArgLeuSerLeu-----GlyAla 370
DB 953 TTTAAGGACGAGAGCTTAAAGAACGAAAGACGTTTAAAGATGTGGAGCAAAAGCTG 1012
QY 371 SerHisGlySerSerValGlnSerLeuProProThrSer----- 383
DB 1013 ACAGAGCAGTCCAAACACCCCGCTTCTTGCTTGCTGCAAGAATGCTGAGAGTCT 1072
QY 384 -----GluGlyLysArgMetSerAlaAspMetSerGluIleGluAlaArgIleAla 401
DB 1073 TACTTTGAATCTTAAACAGAGAGATCAACAGTGCAGAGATGTCATGTGATACACAGCA 1132
QY 402 ThrThr---GlyAsnGlyGlnProArgArgSerIleGlnAspLeuThrValThrGly 420
DB 1133 ACAAGTACGGGAGGCGCATGCGATGCAACCCAAAGTCTCAGCGCATGTGCTGATGGGC 1192
QY 421 ThrGluProGlyGlnValSerSerArgSerSerProSerVal 435
DB 1193 TTT-----GGGATATCTTCTCAATGAACAAAGCAGAGCTGTG 1231
RESULT 12
US-09-221-528-4
/ Sequence 4, Application US/09221528
/ Patent No. 6190874
/ GENERAL INFORMATION:
/ APPLICANT: Acton, Susan
/ TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
/ FILE REFERENCE: MN1-050
/ CURRENT APPLICATION NUMBER: US/09/221,528
/ EARLIER FILING DATE: 1998-12-28
/ EARLIER APPLICATION NUMBER: 09/163,115
/ NUMBER OF SEQ ID NOS: 15
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 4
/ LENGTH: 2120
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (47)..(1411)
US-09-221-528-4
Alignment Scores:
Pred. No.: 4,97e-36 Length: 2120
Score: 490.00 Matches: 138
Percent Similarity: 48.798 Conservative: 84
Best Local Similarity: 30.338 Mismatches: 159
Query Match: 16.264 Indels: 74
DB: 4 Gaps: 19
US-09-830-144-2 (1-579) x US-09-221-528-4 (1-2120)
QY 10 SerSerSerSerSerAlaGlyGluMetIle-----GluAlaProSerGlnVal 25
DB 2 TCGACCCACCGCTCCGCGAAGTAAATCTTGTTCATTATGAGATGTCTCTCCGGT 61
QY 26 LeuAsnPheGluGlnIleAspTyrLysGluIleGluValGluGluValGlyArgGly 45
DB 62 GCCTCCTTGTGCAAAATTAATTGATGACTGCAGTTTGTAAAGACGCGGTGAGGA 121

QY 46 AlaPheGlyValAlaCysLysAlaLysTyr-----ArgAlaLysAspValAlaIleLys 63
DB 122 AGTTTGGAGGTGTATTATCGAGCCAAATGATATACAGACAGAGAGGATGGCTGTAAG 181
QY 64 -----GlnIleGluSerGlnSerGluValArgLysAlaPheIleValGluLeuArgGln 80
DB 182 AAGCTCTCTCAAAATTAAGAAAGAGCGCAAA-----ATA 214
QY 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnPro--- 99
DB 215 CTCAGTCTCTCACTACAGAAACATATCCAGCTTTATGAGTATTTGAACTTCC 274
QY 100 ---ValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGly 118
DB 275 AACTATGCGATTGTCAGAAATATGCTTCTTGAGTCACTATGATTAACATTAACAGT 334
QY 119 -----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerThrCysLeuGln 136
DB 335 AACAGAGTGAAGAGATG-----GATATGATCACTTATTAACCTGGGCCACTGAT 385
QY 137 CysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAsp 156
DB 386 GTAGCCAAAGGATGCATTTTATCATATGAGAGCTCCGTCAAGGTGATTCACAGAGAC 445
QY 157 LeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPhe 176
DB 446 CTCAAATGCAAGAAACGTTTATATAGCTGCTGATGGA---GTACTGAAGATCTGTGACTTT 502
QY 177 GlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys-----GlySerAlaAla 194
DB 503 GGT---GCCCTCGGTTCCATTAACCATACACACAGCTGCTTGATGGAACCTTTCCCA 559
QY 195 TrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSer 214
DB 560 TGTATGCTCTCCAGAAAGTTATCCAGAGCTCCCGTGCAGAAACCTGTGACACATATTC 619
QY 215 TrpGlyIleIleLeuTyrGluValIleThrArgArgLysProPheAspGluIleGlyGly 234
DB 620 TATGTTGGTTCCTCTGGAGATGCTTAACAGAGAGCTCCCTTTAAAGGTGTGGAAGGA 679
QY 235 ProIaPheArgIleMetTyr---AlaValHisAsnGlyThrArgProProLeuIleLys 253
DB 680 -----TTACAAAGTAGCTTGCTGTGATGTGAAAAAACAAGAGATTAACTATCCAAAG 733
QY 254 AsnLeuProLysProIleGluSerLeuMetThrArgCysTyrPheLysAspProSerGln 273
DB 734 AGTTGCCCCAGAAAGTTTGTGGAAGCTGTATCATCAGTGTGGAGAGCTATGCCAAGAAA 793
QY 274 ArgProSerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGly 293
DB 794 CGGCCATCATTCACAGCAAACTATTCATTCCTG----- 826
QY 294 AlaaspGluProLeuGlnTyrProCysGlnTyrSerAspGluGlnSerAspSerAla 313
DB 827 -----GAGTCATGTCAAATGACAG 847
QY 314 ThrSerThrGlySerPheMetAspIleAlaSerThrAsnThrSerAsnLysSerAspThr 333
DB 848 -----AGCCTTCCTGCACAAAGTGAATCACTTCCTACACAAAGCGGAGTGG 895
QY 334 AsnMetGluGlnValProAlaThrAsnAspThrIleLysArgLeuGlnSerLysLeuLeu 353
DB 896 AGGTGGCAA---ATTGAGGCAACTCTTGAGAGGCTAAAGAAACTAGAGCGTGAATTCAGC 952
QY 354 LysAsnGlnAlaLysGlnGlnSerGlnSerGlnArgLeuSerLeu-----GlyAla 370
DB 953 TTTAAGGACGAGAGCTTAAAGAACGAAAGACGTTTAAAGATGTGGAGCAAAAGCTG 1012
QY 371 SerHisGlySerSerValGlnSerLeuProProThrSer----- 383
DB 1013 ACAGAGCAGTCCAAACACCCCGCTTCTTGCTTGCTGCAAGAATGCTGAGAGTCT 1072

QY 384 -----GluGlyLysArgMetSerAlaAspMetSerGluLeuGluAlaAlaAla 401
 Db 1073 TACTTTGAATCTAAACACAGAGAGTCAACAGTGCAGAGATGTCATGTCAGATCACAGCA 1132
 QY 402 ThrThr---GlyAsnGlyGlnProArgArgSerileGlnAspLeuThrValThrGly 420
 Db 1133 ACAAGTAAAGGGAGGCCATGGCATGAACCAAGTCTGCAGGCCATGATGCTGATGGGC 1192
 QY 421 ThrGluProGlyGlnValSerArgSerSerProSerVal 435
 Db 1193 TTT-----GGGGATATCTTCTCAATGAACAAAGCAGGAGCTGTG 1231

RESULT 13
 US-09-593-553-4
 ; Sequence 4, Application US/09593553
 ; Patent No. 6200770
 ; GENERAL INFORMATION:
 ; APPLICANT: Acton, Susan
 ; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
 ; FILE REFERENCE: MNI-050
 ; CURRENT APPLICATION NUMBER: US/09/593,553
 ; CURRENT FILING DATE: 2000-06-14
 ; PRIOR APPLICATION NUMBER: 09/163,115
 ; PRIOR FILING DATE: 1998-09-28
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: Patent in Ver. 2.0
 ; SEQ ID NO 4
 ; LENGTH: 2120
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (47)..(1411)
 US-09-593-553-4

Alignment Scores:
 Pred. No.: 4,97e-36 Length: 2120
 Score: 490.00 Matches: 138
 Percent Similarity: 48.79% Conservative: 84
 Best Local Similarity: 30.33% Mismatches: 159
 Query Match: 16.26% Indels: 74
 DB: Gaps: 19

US-09-830-144-2 (1-579) x US-09-593-553-4 (1-2120)

QY 10 SerSerSerSerAlaGlyGluMetIle-----GluAlaProSerGlnVal 25
 Db 2 TCACCCACGCGTCCGGTGAAGTATAAATCTTGTCTATGAGATGCTCTCGGT 61
 QY 26 LeuAsnGluGluLeuAspTyrLysGluLeuGluValValGlyArgGly 45
 Db 62 GCCTCCTTTGTGCAAAATTAATTTGATGACTTGACGTTTTTTGAAACTGCGGTGAGGA 121
 QY 46 AlaPheGlyValValCysLysAlaLysTyr-----ArgAlaLysAspValAlaLys 63
 Db 122 AGTTTGGGAGTGTATTCGAGCAATATGATATCATCAGCAAGAGGAGTGTCTGAAG 181
 QY 64 -----GlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
 Db 182 AAGCTCTCAAAATAGAAAGAGGAGCA-----ATA 214
 QY 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnPro--- 99
 Db 215 CTCAGTGTCTCAGTCACAGAAACATCATCCAGTTTTATGGAGTAATCTTGAACCTCCC 274
 QY 100 ---ValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGly 118
 Db 275 AACTATGGCATGTCTCAGAAATATGCTTCTGGGATCACTCTATGATTACATTAAACAGT 334
 QY 119 -----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeuGln 136
 Db 335 AACAGAAGTGAGGAGATG-----GATATGGATCACATTATACCTGGGCCACTGAT 385

QY 137 CysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAsp 156
 Db 386 GTAGCAAAAGGATCATATTTACATATGAGGCTCCTGTCAAGGTGATTCACAGAGAC 445
 QY 157 LeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPhe 176
 Db 446 CTCAAAGTCAAGAAAGCTTGTATAGTCTCTGATGGA---GTACTGAAGATCTGTGACTTT 502
 QY 177 GlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys-----GlySerAlaAla 194
 Db 503 GGT---GCCTCTCGGTTCCTCAATCAACACACATGCTCTTGGTTGGAACCTTCCCA 559
 QY 195 TrpMetAlaProGluValPheGluLysSerAsnTyrSerGluLysCysAspValPheSer 214
 Db 560 TGGATGGCTCCAGAGATTATCCAGAGTCTCCTGTGTGAGAACTTGTGACACATATTC 619
 QY 215 TrpGlyIleLeuLeuTyrGluValIleThrArgArgLysPropheAspGluIleGly 234
 Db 620 TATGTTGTGTTCTCTGGGAGATGCTAAAGGAGGAGTCCCTTTAAAGGTTTGAAGGA 679
 QY 235 ProAlaPheArgIleMetTyr---AlaValHisAsnGlyThrArgProProLeuIleLys 253
 Db 680 -----TTACAAGTAGCTTGGCTTGTAGTGAAAAAACGAGAGATTAAACCATCCAAGC 733
 QY 254 AsnLeuProLysProLleGluSerLeuMetThrArgCysTyrSerLysAspProSerGln 273
 Db 734 AGTTGCCCCAGAGTTTGTGTAAGTGTATCATCAGTGTGGGAGGCTGATGCCAAGAA 793
 QY 274 ArgProSerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGly 293
 Db 794 CGGCCATCATTAAGCAAAATCATTTCAATCTCTG----- 826
 QY 294 AlaAspGluProLeuGlnTyrProCysGlnTyrSerAspGluGlyGlnSerAsnSerAla 313
 Db 827 -----GAGTCCATGTCAAAATGACACG 847
 QY 314 ThrSerThrGlySerPheMetAlaSerThrAsnThrSerAsnLysSerAspThr 333
 Db 848 -----AGCCTTCCTGACAAAGTGAACCTCATCTCTACACAACAGGCGGAGTGG 895
 QY 334 AsnMetGluGlnValProAlaThrAsnAspThrIleLysArgLeuGluSerLysLeuLeu 353
 Db 896 AGGTGCGAAA---ATTGAGGCAACTCTTCAGAGGCTAAAGAAACTAGAGCGTGTATCTCAGC 952
 QY 354 LysAsnGlnAlaLysGlnSerGluSerCysArgLeuSerLeu-----GlyAla 370
 Db 953 TTTAAGCAGCAGGAGCTTAAAGAACGAGAACGCTTTAAAGATGTGGGAGCAAAAGCTG 1012
 QY 371 SerHisGlySerSerValGluSerLeuProProThrSer----- 383
 Db 1013 ACAGAGCAGTCCACACACCCCGCTCTCTTGCCTCTGCTGCAAGAATGCTGAGGAGTCT 1072
 QY 384 -----GluGlyLysArgMetSerAlaAspMetSerGluIleGluAlaAlaAla 401
 Db 1073 TACTTTGAATCTAAACACAGAGGAGTCAACAGTGCAGAGATGTCATGTCAGATCACAGCA 1132
 QY 402 ThrThr---GlyAsnGlyGlnProArgArgSerIleGlnAspLeuThrValThrGly 420
 Db 1133 ACAAGTAAAGGGAGGCCATGGCATGAACCAAGTCTGCAGGCCATGATGCTGATGGGC 1192
 QY 421 ThrGluProGlyGlnValSerArgSerSerProSerVal 435
 Db 1193 TTT-----GGGGATATCTTCTCAATGAACAAAGCAGGAGCTGTG 1231

RESULT 14

US-09-221-237-4
 ; Sequence 4, Application US/09221237
 ; Patent No. 6214597
 ; GENERAL INFORMATION:
 ; APPLICANT: Acton, Susan
 ; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
 ; FILE REFERENCE: MNI-050
 ; CURRENT APPLICATION NUMBER: US/09/221,237

CURRENT FILING DATE: 1998-12-28
 EARLIER APPLICATION NUMBER: 09/163,115
 EARLIER FILING DATE: 1998-09-29
 NUMBER OF SEQ ID NOS: 15
 SOFTWARE: Patent In Ver. 2.0
 SEQ ID NO 4
 LENGTH: 2120
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (47)..(1411)
 US-09-221-237-4

Alignment Scores:

Pred. No.:	4,97e-36	Length:	2120
Score:	490.00	Matches:	138
Percent Similarity:	48.79%	Conservative:	84
Best Local Similarity:	30.33%	Mismatches:	159
Query Match:	16.26%	Indels:	74
DB:	4	Gaps:	19

US-09-830-144-2 (1-579) x US-09-221-237-4 (1-2120)

10 SerSerSerSerSerAlaGlyGluMetIle-----GluAlaProSerGlnVal 25
 2 TCGACCCAGCGCGTGGAGATTAATACCTTGTTCATTATGATGTCGCTCGGT 61
 26 LeuAsnPhcGluGluIleAspTyrIleGluIleGluValGluValValGlyArgGly 45
 62 GCCTCCTTGTGCAAAATTAAATTGATGCTGACGATTTTGGAAACGCGGTGAGGA 121
 46 AlaPhcGluValValCysLysAlaLysTrp-----ArgAlaLysAspValAlaIleLys 63
 122 AGTTTGGAGGCTTATATGAGCCAAATGATATCAGACAGACAGAGGTGCTGTAAG 181
 64 -----GlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
 182 AAGCTCCTCAAAATAGAGAAAGAGCAGAA-----ATA 214
 81 LeuSerArgValAlaAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnPro--- 99
 215 CTCAGAGTCTCTGACACAGAAACATCATCATGTTTATGAGGATTAATTTTGAACCTCC 274
 100 ---ValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGly 118
 275 AACTATGCAATTCACAGAAATATGCTTCTGCGGATCATCTATGATATACATTAAAGT 334
 119 -----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeuGln 136
 335 AACAGAAAGTGAAGAGATG-----GATATGATCATCATTAATGACCTGGGCCACTGAT 385
 137 CysSerGlnGlyAlaAlaTyrIleLeuHisSerMetGlnProLysAlaLeuIleHisArgAsp 156
 386 GTAGCCAAAGGAATGCAATTATTAATGATGAGAGCTCTCTGTCAGAGTGTACACAGAGAC 445
 157 LeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPhe 176
 446 CTCAGAGTCAAGAAACGTTGTTATAGCTGCTGATGGA---GTCATGAGATCTGTGCTTT 502
 177 GlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys-----GlySerAlaAla 194
 503 GGT---GCCCTCGGTTCCATACACATACACACATGCTCTTGTTGGAACCTTCCCA 559
 195 TyrMetAlaProGluValAlaPheGluGlySerAsnTyrSerGluLysCysAspValPheSer 214
 560 TGGATGGCTCCAGAAAGTTATCCAGAGTCTCTGCTGTCGAAACCTTGTAACATATTTCC 619
 215 TyrGlyIleIleLeuTyrGluValAlaIleThrArgArgLysProPheAspGluIleGlyGly 234
 620 TATGCTGTGTTCTCTGAGAGATGCTAAACAGGAGGTCCCTTTAAAGTTTGAAGAAG 679
 235 ProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgProProLeuIleLys 253

680 -----TTACAGTACTGCTTGCTTAGTGAAAAAAGCAGAGATTAACTTCACAGC 733
 254 AsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGln 273
 734 AGTTGCCCAAGAAAGTTTGTCTGAACGTGTACATCAGTGTGGAGAGCGATCCCAAGAA 793
 274 ArgProSerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGly 293
 794 CGGCCATCATTCACAGCAAAATCATTTCAATCCCTG----- 826
 294 AlaAspGluProLeuIleTyrProCysGlnTyrSerAspGluGlyGlnSerAsnSerAla 313
 827 -----GAGTTCATGTCATTAATACAGC 847
 314 ThrSerThrGlySerPheMetAspIleAlaSerThrAsnThrSerAsnLysSerAspThr 333
 848 -----AGCTTCTCTGACAGTGAATCTTCTTCAACAAACAGCGGAGTGC 895
 334 AsnMetGluGlnValProAlaThrAsnAspThrIleLysArgLeuGluSerLysLeuLeu 353
 896 AGGTGCGAA---ATTGAGGCAACTCTTGAGAGGCTTAAAGAAACTGAGCGTATCTCAGC 952
 354 LysAsnGlnAlaLysGlnGlnSerGluSerGlyArgLeuSerLeu-----GlyAla 370
 953 TTTAAGGAGCAGAGCTTAAAGAACAGAAAGACGTTTAAAGTGTGGAGCAAAAGCTG 1012
 371 SerHisGlySerSerValGluSerLeuProProThrSer----- 383
 1013 ACAGAGAGCTCAACACCCGCTTCTTGCCTTGTGCTGCAAGATGTCTAGAGAGTCT 1072
 384 -----GluGlyLysArgMetSerAlaAspMetSerGluIleGluAlaArgIleAlaAla 401
 1073 TACTTGAATCTTAAACAGAGAGATCAACAGTCAAGATGATGATGATGATGATGATGATGAT 1132
 402 ThrThr---GlyAsnGlyGlnProArgArgSerIleGlnAspLeuThrValThrGly 420
 1133 ACAAGTACCGGAGGCGCCATGAGATGACCAAGAGTCTGAGCCATGATGCTGATGAGGC 1192
 421 ThrGluProGluGlnValSerSerArgSerSerSerProSerVal 435
 1193 TTT-----GGGATATCTTCTCAATGAAACAAAGCAGAGCTGTG 1231
 RESULT 15
 US-09-221-235-6
 Sequence 6, Application US/09221235
 Patent No. 6043040
 GENERAL INFORMATION:
 APPLICANT: Acton, Susan
 TITLE OR INVENTION: NOVEL CSARK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
 FILE REFERENCE: NMT-050
 CURRENT APPLICATION NUMBER: US/09/221,235
 EARLIER FILING DATE: 1998-12-28
 EARLIER APPLICATION NUMBER: 09/163,115
 NUMBER OF SEQ ID NOS: 15
 SOFTWARE: Patent In Ver. 2.0
 SEQ ID NO 6
 LENGTH: 1365
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (1)..(1365)
 US-09-221-235-6

Alignment Scores:

Pred. No.:	1.69e-35	Length:	1365
Score:	481.00	Matches:	132
Percent Similarity:	49.08%	Conservative:	81
Best Local Similarity:	30.41%	Mismatches:	151
Query Match:	15.96%	Indels:	70
DB:	3	Gaps:	18

US-09-830-144-2 (1-579) x US-09-221-235-6 (1-1365)

```
QY 27 AsnPheGluGluIleAspTyrLysGluIleValGluValValGlyArgGlyAla 46
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 19 TCCTTTGTGCAATTAATTTGATGCTTCAGATTTTGTGAAACATCGCGTGGAGAA 78
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 47 PheGlyValValCysLysAlaLysTrp-----ArgAlaLysAspValAlaIleLys--- 63
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 79 TTTGGAGTGTATATCAGAGCAATATGATATCACAGGCAAGGAGGTGGCTGTAAGAAG 138
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 64 -----GlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGlnLeu 81
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 139 CTCCTCAAAATAGAGAAAGAGGCGAA-----ATACTC 171
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 82 SerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnPro----- 99
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 172 AGTGTCTCAGTCACAGAAACATCATCCAGTTTATGGAGTAATTTCTGNAACCTCCCAAC 231
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 100 ValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGly--- 118
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 232 TATGGCATTTGCAGAAATATGCTCTCTGGGATCACTCTATGATTACATTAAACAGTAAC 291
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 119 ---AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeuGlnCys 137
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 292 AGAAGTGAGGAGATG-----GATATGGATCACATTATGACCTGGGCCCATGATGTA 342
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 138 SerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeu 157
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 343 GCCAAGGAATGCATTATTACATATGGAGGTCCTGTCAAGGTGATTACAGAGACCTC 402
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 158 LysProProAsnLeuLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGly 177
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 403 AAGTCAAGAAACGTTGTTATAGCTGTGATGGA---GTACTGAAGATCTGTGACTTTGT 459
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 178 ThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys-----GlySerAlaAlaTrp 195
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 460 ---GCCTCTCGGTCCATAACCATCAACACACACATGCTCTGGTTGGAACTTTCCCATGG 516
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 196 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 215
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 517 ATGGCTCCAGAAGTTATCCAGAGTCCCTGTGTGAGAACTTTGTGACACATATTCCTAT 576
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 216 GlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 235
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 577 GGTGTGGTCTCTGGGAGATGCTAAACAAGGAGGTCCTCTTAAAGTTTGGAAAGA--- 633
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 236 AlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgProProLeuIleLysAsn 254
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 634 ---TTACAAGTAGCTTGGCTTGTAGTGGAAAAAACAAGAGAGATTAAACCATCCAAAGCAGT 690
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 255 LeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArg 274
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 691 TGCCCCCAGAAGTTTTCCTGAACCTGTTACATCAGTGTGTGGAGACTGATGCCAAGAAACGG 750
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 275 ProSerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAla 294
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 751 CCATCATTCAGCAAAATCATTTCAATCTCTG----- 780
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 295 AspGluProLeuGlnTyrProCysGlnTyrSerAspGluGlyGlnSerAsnSerAlaThr 314
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 781 -----GAGTCCATGTCAAAATGACACG--- 801
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 315 SerThrGlySerPheMetAspIleAlaSerThrAsnThrSerAsnLysSerAspThrAsn 334
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 802 -----AGCTTCCTGCACAAGTGAATCACTTCCTCACACAAACAAGCGGAGTGGAGG 852
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 335 MetGluGlnValProAlaThrAsnAspThrIleLysArgLeuGluSerLysLeuLys 354
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 853 TCGGAA---ATTGAGGCAACTCTTGAGAGGCTAAAGAAACTAGAGCGTGATCTCAGCTTT 909
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 355 AsnGlnAlaLysGlnSerGluSerGlyArgLeuSerLeu-----GlyAlaSer 371
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
Db 910 AAGGAGCAGGAGCTTTAAAGAACGAGAAAGACGTTTAAAGATGTGGGAGCAAAAGCTGACA 969
QY 372 HisGlySerSerValGluSerLeuProProThrSer----- 383
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 970 GAGCAGTCCAAACACCCCGCTTCTCTTGGCTCTTGTGCAAGAATGCTCTGAGGAGTCTTTAC 1029
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 384 ---GluGlyLysArgMetSerAlaAspMetSerGluIleGluAlaArgIleAlaAlaThr 402
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1030 TTTGAAATCTAAACACAGAGGAGTCAAAACAGTGCAGAGATGTCTGTCTCAGATCACAGCAACA 1089
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 403 Thr---GlyAsnGlyGlnProArgArgSerIleGlnAspLeuThrValThrGlyThr 421
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1090 AGTAAACGGGGAGGCCATGGCATGAACCCAAAGTCTGCAGGCCATGATGCTGATCGGCTTT 1149
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 422 GluProGlyGlnValSerSerArgSerSerProSerVal 435
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1150 -----GGGGATATCTTCTCAATGAACAAACAGCAGGAGCTGTG 1185
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: December 10, 2002, 02:06:18
Job time : 92.482 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 9, 2002, 23:04:29 ; Search time 83.4017 Seconds
(without alignments)
2707.033 Million cell updates/sec

Title: US-09-830-144-2

Perfect score: 3014

Sequence: 1 MSTAASASSSSSSAGEMIE.....QCKKOLEVIRSQQKROGTS 579

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 350425 seqs, 194966369 residues

Total number of hits satisfying chosen parameters: 700850

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09830144/runat_04122002_141354_2307/app_query.fasta_1.1422
-DB=Published_Applications_NA -QFWT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=Blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09830144 @CGN 1.1 51.0runat_04122002_141354_2307
-NCPUS=6 -ICPU=3 -NO_XLPHY -NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DRV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3014	100.0	1788	9	US-10-158-895-14
2	3014	100.0	2656	9	US-10-158-895-3
3	493.5	16.4	3454	10	US-09-969-347-226
4	490	16.3	2120	10	US-09-757-982-4

ALIGNMENTS

RESULT 1

US-10-158-895-14

; Sequence 14, Application US/10158895

; Patent No. US20020155624A1

; GENERAL INFORMATION:

; APPLICANT: ONO, KOICHIRO

; APPLICANT: OHTOMO, TOSHIHIKO

; APPLICANT: TSUCHIYA, MASAYUKI

; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES

; FILE REFERENCE: 053466/0278

; CURRENT APPLICATION NUMBER: US/10158,895

; CURRENT FILING DATE: 2002-06-03

; PRIOR APPLICATION NUMBER: US/09529,279

; PRIOR FILING DATE: 2000-04-11

; PRIOR APPLICATION NUMBER: PCT/JP98/04796

; PRIOR FILING DATE: 1998-10-22

; PRIOR APPLICATION NUMBER: JP 9/290188

; PRIOR FILING DATE: 1997-10-22

; NUMBER OF SEQ ID NOS: 48

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 14

; LENGTH: 1788

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (7)..(1776)

US-10-158-895-14

Sequence 6, Appli
Sequence 1, Appli
Sequence 3, Appli
Sequence 1, Appli
Sequence 3, Appli
Sequence 7, Appli
Sequence 9, Appli
Sequence 1014, Ap
Sequence 1, Appli
Sequence 1577, Ap
Sequence 1073, Ap
Sequence 882, App
Sequence 16, Appl
Sequence 903, App
Sequence 84, Appl
Sequence 405, App
Sequence 2, Appli
Sequence 2, Appli
Sequence 48, Appl
Sequence 271, App
Sequence 1, Appli
Sequence 3, Appli
Sequence 1085, Ap
Sequence 1, Appli
Sequence 1405, Ap
Sequence 491, App
Sequence 4, Appli
Sequence 23, Appl
Sequence 1, Appli
Sequence 207, App
Sequence 1, Appli
Sequence 1, Appli
Sequence 3, Appli
Sequence 553, App
Sequence 1, Appli
Sequence 1479, Ap
Sequence 1128, Ap
Sequence 34, Appl
Sequence 7, Appli
Sequence 13, Appli

Alignment Scores:

Pred. No.:	5,93e-251	Length:	1788
Score:	3014.00	Matches:	579
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-830-144-2 (1-579) x US-10-158-895-14 (1-1788)

```

QY 1 MetSerThrAlaSerAlaIaSerSerSerSerSerSerAlaGluMetIleGlu 20
Db 7 ATGCTCAAGCCCTCCGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCG 66
QY 21 AlaProSerGlnValLeuAnpHeGluGluIleAspTyrLeuGluIleGluVal 40
Db 67 GCCCTTCCAGGTCTCTCAACTTGAAGAGATGACTACAGAGAGATCGAGGTGGAAG 126
QY 41 ValValGlyArgGlyAlaPheGlyValValCysLeuAlaLysTyrAlaLysAspVal 60
Db 127 GTTGTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 186
QY 61 AlaIleLeuSerGlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
Db 187 GCTATTAAACAAATAGAAATGAAATCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 246
QY 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyValCysLeuAnpProVal 100
Db 247 TTATCCCGTGTGAACCATCTTAATATTGTAAACCTTTATGAGAGAGAGAGAGAG 306
QY 101 CysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnValLeuHisGlyAlaGlu 120
Db 307 TGTCTGTGATGAAATATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 366
QY 121 ProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeuGlnCysSerGlnGly 140
Db 367 CCATTCGCAATTAATATCTGCTGCCACGCAATAGTGGTGTGTACAGAGAGAGAG 426
QY 141 ValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysProPro 160
Db 427 GTGGCTTATCTTACAGCAACGCAACCCAAAGCCCTAATTCACAGAGAGAGAGAG 486
QY 161 AsnLeuLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCys 180
Db 487 AACTTACTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 546
QY 181 AspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyrMetAlaProGluVal 200
Db 547 GACATTCAGACACACATGACCAATTAACAGAGAGAGAGAGAGAGAGAGAGAGAG 606
QY 201 PheGluGlySerAsnTyrSerGluLysCysAspValPheSerTyrGlyIleIleLeuTyr 220
Db 607 TTGGAAGGTGATTAATCACTGAAATAATGAGAGAGAGAGAGAGAGAGAGAGAG 666
QY 221 GluValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMet 240
Db 667 GAAGTGATTAACGGCGTGGAAACCTTTGATGAGATGAGAGAGAGAGAGAGAGAG 726
QY 241 TyrAlaAlaHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGlu 260
Db 727 TGGGCTGTATATATGATCTGACACCACTATATAAAATTTACCTAAGGCCATTGAG 786
QY 261 SerLeuMetThrArgCysTyrSerLysAspProSerGlnArgProSerMetGluGluIle 280
Db 787 AGCTTGATGATCGTCTTGTCTTAAAGATCCTTCCAGAGGCCCTTCAATGAGGAAAT 846
QY 281 ValLysIleMetThrHisLeuMetArgTyrPheProGlyValAlaAspGluProLeuGlnTyr 300
Db 847 GTGAATAATATGACTACTGATGCGGTACTTCCAGAGACAGATGAGCCATTACAGTAT 906
QY 301 ProCysGlnTyrSerArgGluGlyGlnSerAsnSerAlaThrSerThrGlySerPheMet 320

```

```

Db 907 CCTGTGCTAGTATTGATGAGAGACAGAGCACTCTGCCACGATACAGGCTCATTCATG 966
QY 321 AspIleAlaSerThrAsnThrSerAsnLysSerAspThrAsnMetGluGluValProAla 340
Db 967 GACATTCCTCTTACAAATATACAGATTAACAAAGATGACACTAATATGAGAGCAAGTCTCGCC 1026
QY 341 ThrAsnAspThrIleLysAsnGluGluSerLysLeuLeuLysAsnGlnAlaLysGlnGln 360
Db 1027 ACAATATATCTATTAAAGCCCTTAGAATCAAAATGTGTGAAATAATGAGCAAGCAACAG 1086
QY 361 SerGluSerGlyArgLeuSerLeuGlyAlaSerHisGlySerSerValGluSerLeuPro 380
Db 1087 AGTGAATCTGAGCCTTTAAGCTTGGAGAGCTCCCATGAGAGAGAGAGAGAGAGAGAGAG 1146
QY 381 ProThrSerGluGlyLysArgMetSerAlaAspMetSerGluIleGluAlaArgIleAla 400
Db 1147 CCAACTCTGAGGCGCAAGAGAGATGAGTGCATGCAATGCTGTAATAATAGAGCTAGAGATCGCC 1206
QY 401 AlaThrThrGlyAsnGlyGlnProArgArgArgSerIleGlnAspLeuThrValThrGly 420
Db 1207 GCAACCAAGGCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1266
QY 421 ThrGluProGlyGlnValSerSerArgSerSerSerProSerValArgMetIleThrThr 440
Db 1267 ACAGAACTGTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1326
QY 441 SerGlyProThrSerGluLysProThrArgSerHisProTyrThrProAspAspSerThr 460
Db 1327 TCAGAGACCAACTCAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1386
QY 461 AspThrAsnGlySerAspAsnSerIleProMetAlaTyrLeuThrLeuAspHisGlnLeu 480
Db 1387 GATACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1446
QY 481 GlnProLeuAlaProCysProAsnSerLysGluSerMetAlaValPheGluGlnHisCys 500
Db 1447 CAGCCTCTAGACAGCGTCCCAACTCCAAAGAGATCTATGAGAGAGAGAGAGAGAGAG 1506
QY 501 LysMetAlaGlnGluTyrMetLysValGlnThrGluIleAlaLeuLeuLeuGlnArgLys 520
Db 1507 AAAATGGCACAAGATATATGAAAGTTCAAAACAGAAATTCATTGTTATATACAGAGAA 1566
QY 521 GlnGluLeuValAlaGluLeuAspGlnAspGluLysAspGlnGlnAsnThrSerArgLeu 540
Db 1567 CAAAGACTAGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1626
QY 541 ValGlnGluHisLysLysLeuLeuAspGluAsnLysSerLeuSerThrTyrTyrGlnGln 560
Db 1627 GTACAGGAACATTAAGAGCTTTTATGATGAAACAAAGAGAGAGAGAGAGAGAGAG 1686
QY 561 CysLysLysGlnLeuGluValIleArgSerGlnGlnGlnLysArgGlnGlnLysSer 579
Db 1687 TGCAAAAACAACTAGAGAGTATCAGAAATCAGAGAGAGAGAGAGAGAGAGAGAGAG 1743

```

RESULT 2

```

US-10-158-895-3
; Sequence 3, Application US/10158895
; Patent No. US20020155624A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF- $\beta$  INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48

```



```

CURRENT FILING DATE: 2001-10-02
PRIOR APPLICATION NUMBER: US/60/237,598
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: US/60/237,604
PRIOR FILING DATE: 2000-10-03
NUMBER OF SEQ ID NOS: 318
SOFTWARE: PatentIn version 3.0
SEQ ID NO 226
LENGTH: 3454
TYPE: DNA
ORGANISM: Homo sapiens
US-09-969-347-226

Alignment Scores:
Pred. No.: 2,75e-33 Length: 3454
Score: 493.50 Matches: 145
Percent Similarity: 46.46% Conservative: 87
Best Local Similarity: 29.06% Mismatches: 191
Query Match: 16,37% Indels: 79
DB: 10 Gaps: 14

US-09-830-144-2 (1-579) x US-09-969-347-226 (1-3454)
QY 21 AlaProSerGlnValLeuAsnProHepGluIleAspTyrLeuGluIleGluValGluGlu 40
DB 535 GCACCCGCGGGCTCCAGCTGCCCCAGAGATCCCTTCCAGAGCTGAGAGAG 594
QY 41 ValValGlyArgGlyAlaPheGlyValAlaCysLeuAlaLysTyrArgAlaLysAspVal 60
DB 595 ATCATCGTGTGGGGGCTTTGGCAAGGTCTATCGGGCCCTGTGGCGGAGAGGTG 654
QY 61 AlaIleLeuGlnIleGluSerGluSerGluValArgLysAlaPheIleVal----- 76
DB 655 GCAGTCAAGGCCCCCGGCTGAGACCTCGAGAGAGACCCGAGTGCACGAGAGAGTG 714
QY 77 -----GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGly 94
DB 715 TGGCAGAAAGCCCGGCTTTGGAGCCCTGACAGACCCCAACATATGCTTACGGGCG 774
QY 95 AlaCysLeuAsnPro-----ValCysLeuValMetGluTyrAlaGluIleGlySerLeu 112
DB 775 GCGTGGCTCAACCCCGCACACCTCTGCTAGTATGAGATGCCCGGGGTGGTGCACCTG 834
QY 113 TyrAsnValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSer 132
DB 835 AGCAGAGTGTGGCAGGTGCCCGGGTGCACCTCACGTG-----CTGGTCAAC 882
QY 133 TyrCysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeu 152
DB 883 TGGGCTGTGAGGTGGCCCGGGGATGAATCTACTACCAATGATGCCCTGTGCCCATC 942
QY 153 IleHisArgAspLeuLysProProAsnLeuLeuLeuVal----- 165
DB 943 ATCCACCGGAGACTCAAGTCCATCAACATCTCTGAGGCCATCGAGAACCAACAC 1002
QY 166 AlaGlyLeuIleValLeuLysIleCysAspPheGlyThrAlaCysAsp---IleGlnThr 184
DB 1003 CTCGCAGACACGAGTGTCAAGATCACGAGCTTGGCCCTGCCCGCAGTGGACACAGACC 1062
QY 185 HisMetThrAsnAsnLysSerAlaAlaTyrMetAlaProLysValPheGluGlySer 204
DB 1063 ACCAAGATGAGCGCTGGGAGACTTACGCTGTGATGGCGGAGGTATTCGCTCTTCC 1122
QY 205 AsnTyrSerGluLysCysAspValPheSerTyrGlyIleIleLeuTyrGluValIleThr 224
DB 1123 CTCTTCTCCAAAGACTGATGCTGTGAGGCTTGGGGGTGCTGTGGAGAGCTGCTGAGC 1182
QY 225 ArgArgLysProPheAspGluIleGlyLysProAlaPheArgIleMetTyrAlaVal--- 243
DB 1183 GGGAGAGTCCCTTACCGTGAATC-----GACGCTTGGCGTGGCGATATGGCGTGGCT 1236
QY 244 HisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGluSerLeuMet 263

```

```

DB 1237 ATGAATAAGCTGACGGCTGCCCATTTCCCTCCACGTGCCCGAGGCCCTTTGGCCGCTCTG 1296
QY 264 ThrArgCysTyrPheSerLysAspProSerGlnArgProSerMetGluIleValLysIle 283
DB 1297 GAGGAATGCTGGGACCCAGACCCCGAGGCGGCGAGATTTCCGTAGCATCTTGAACCG 1356
QY 284 MetThrHisLeuMetArgTyrPhePro----- 292
DB 1357 CTTAAGTCAATCAACAGTC-AGCCTGTTCAGATGCCACTGAGTCTTCACATCGCT 1415
QY 292 ----- 292
DB 1416 GCAGAAAGACTGGAAGCTGGAATTCAGACATGTTGATGACCTTGCAGCAAGAGAA 1475
QY 293 GlyAlaAspGluProLeuGlnTyrProCysGlnTyrSerAspGluIleGlnSerAsnSer 312
DB 1476 GCAGCTTCGGGCGCGTG-AGGAGAGCTGTGGCGGCGGACAGAGACAGCGCTTCCAGG 1534
QY 313 AlaThrSerThrGly-----SerPheMetAspIleAlaSerThrAsnThrSer 328
DB 1535 AGGACAGCTGCGCGCGGCGGAGCAGAGCTGGCAGAACTGAGATGACATCTGGAAC 1594
QY 329 AsnLysSerAspThrAsnMetGluGlnValProAlaThrAsnAspThrIleLysArgLeu 348
DB 1595 GCGAGCTGCACCTGCTC-ATGTGCAAGCTGAGCCAGAGAGAGCCCGGCTCGCAAGCGC 1653
QY 349 GluSerLysLeuLysAsnGlnAlaLysGlnGlnSerGluSerGly-----Arg 365
DB 1654 AAGGGCAACTTCAAGCCACACCGCTGCTCAAGCTGCGGAGAGCGGACGCCACATCAGC 1713
QY 366 LeuSerLeuGlnAlaSerHisGlySerSerValGluSerLeuProProThrSerGluGly 385
DB 1714 CTGCCCTCTGGCTTTGACATTAAGATCAAGTCCAGGCTCTTCAACTCTGAT----- 1767
QY 386 LysArgMetSerAlaAspMetSerGluIleGluAlaArgIleAlaIleThrThrGlyAsn 405
DB 1768 AAGCGAAAGATTCGATGGGGCCAGCCCGCTGCAAGCCCGACATCATC----- 1818
QY 406 GlyGlnProArgArgArgSerIleGlnAspLeuThrValThrGlyThrGluProGlyGln 425
DB 1819 -----CCCGCGCTGAGAGGCGCATTCGC-----CTGACTCCCGTGCATGTGGTGC 1863
QY 426 ValSerSerArgSerSerProSerValArgMetIleThrHisSerGlyProThrSer 445
DB 1864 AGCAGAGTGCAGCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1923
QY 446 GluLys-----ProThrArgSerHisProTyrThrProAspAspSer 459
DB 1924 AAGGAAGAATGTCGGGGGCGCAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980

RESULT 4
US-09-757-982-4
Sequence 4, Application US/09757982
Patent No. US20020094559A1
GENERAL INFORMATION:
APPLICANT: Acton, Susan
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
FILE REFERENCE: NMI-050
CURRENT APPLICATION NUMBER: US/09/757,982
CURRENT FILING DATE: 2001-01-10
PRIOR APPLICATION NUMBER: 09/163,115
PRIOR FILING DATE: 1998-09-29
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 2120
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (47)..(1411)
US-09-757-982-4

```

Alignment Scores:

Pred. No.: 2,84e-33 Length: 2120
 Score: 490.00 Matches: 138
 Percent Similarity: 48.79% Conservative: 84
 Best Local Similarity: 30.33% Mismatches: 159
 Query Match: 16.26% Indels: 74
 DB: 10 Gaps: 19

US-09-830-144-2 (1-579) x US-09-757-982-4 (1-2120)

```

QY 10 SerSerSerSerAlaGlyGluMetIle-----GluAlaProSerGlnVal 25
Db 2 TCAGCCACCGCGTCCGGTGAAGTAACTTTGTCAATTATGATGATCGTCTCGGT 61
QY 26 LeuAsnPhedGluIleAspTyrLysGluIleGluValGluValValGlyArgGly 45
Db 62 GCCTCCTTGTGCAATTAATTTGATGACTTCAGTCTTGTGAACTCGCGTGAAGA 121
QY 46 AlaPheGlyValValCysLysAlaLysTrp-----ArgAlaLysAspValAlaIleLys 63
Db 122 AGTTTGGGAGTGTATTCAGGCCAATGATATCACAGGACAAGGAGGTGGTGTAAAG 181
QY 64 -----GlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
Db 182 AAGCTCTCAAAATAGAGAAGAGGCAGAA-----ATA 214
QY 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnPro--- 99
Db 215 CTAGTGTCTCAGTCACAAACATCATCCAGTTTATGGAGTAATTCTTGAACCTCCC 274
QY 100 ---ValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGly 118
Db 275 AACTATGGCATTGTTCAGAGAATATGCTTCTCTGGATCACTCTATGATTACATTACAGT 334
QY 119 -----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeuGln 136
Db 335 AACAGAAAGTGAGGAGATG-----GATATGGATCACATTATGACCTGGGCCACTGAT 385
QY 137 CysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAsp 156
Db 386 GTAGCCAAAGAAATGCAATTATTACATATGATGAGGCTCCTCTCAAGGTGATTTCAGAGAC 445
QY 157 LeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPhe 176
Db 446 CTCAAGTCAAGAAACGTTGTATAGTGCTGATGGA---GTACTGAAGATCTGTGACTTT 502
QY 177 GlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys-----GlySerAlaAla 194
Db 503 GGT---GCCTCTCGGTTCCATAACCAACATACACACATGTCCTTGTGTGAACCTTCCCA 559
QY 195 TrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSer 214
Db 560 TGGATGCTCCAGAAGTTATCCAGAGTCTCCGTGTGCAGAACTTGTGACATATATCC 619
QY 215 TrpGlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGly 234
Db 620 TATGTTGGTGTCTCTGGAGATGCTAACAGGAGGTCCCTTTTAAAGTTTGGGAAGA 679
QY 235 ProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgProProLeuIleLys 253
Db 680 -----TTACAAGTAGCTTGGCTTGTAGTGGAATAAACCGAGAGATTAACCATTCGAAGC 733
QY 254 AsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGln 273
Db 734 AGTTGCCCAAGAGTTTGTGTAACCTGTTTACATCAGTGTGGGAAGCTGATGCCAAGAA 793
QY 274 ArgProSerMetGluIleValLysIleMetThrHisLeuMetArgTyrPheProGly 293
Db 794 CGGCCATCATCAAGCAATCATTTCAATCTCG-----826
QY 294 AlaAspGluProLeuGlnTyrProCysGlnTyrSerAspGluGlyGlnSerAsnSerAla 313
Db 827 -----GAGTCCATGTCAATGACACG 847

```

```

QY 314 ThrSerThrGlySerPheMetAspIleAlaSerThrAsnThrSerAsnLysSerAspThr 333
Db 848 -----AGCCTTCCTGCAAGAGTGAACCTCATCTCTACACAACAGCGGAGTGG 895
QY 334 AsnMetGluGlnValProAlaThrAsnAspThrIleLysArgLeuGluSerLysLeuLeu 353
Db 896 AGGTGCGGAA--ATTGAGGCAACTCTTGAGAGGCTTAAAGAAACTAGAGCGTGATCTCAGC 952
QY 354 LysAsnGlnAlaLysGlnSerGluSerGlyArgLeuSerLeu-----GlyAla 370
Db 953 TTTAAGGACGAGGAGCTTTAAAGAACGAGAACGAGCTTTAAAGATGTGGGAGCAAAAGCTG 1012
QY 371 SerHisGlySerSerValGluSerLeuProThrSer-----383
Db 1013 ACAGAGCAGTCCAACACACCCCGCTTCTCTTGCTCTTGCTGCAAGAAATGCTGAGGAGTCT 1072
QY 384 -----GluGlyLysArgMetSerAlaAspMetSerGluIleGluAlaArgIleAlaAla 401
Db 1073 TACTTTGAATCTAAACACAGAGGAGTCAAAACAGTGCAGAGATGTCTATGTCAGATCACAGCA 1132
QY 402 ThrThr---GlyAsnGlyGlnProArgArgArgSerIleGlnAspLeuThrValThrGly 420
Db 1133 ACAAGTAACGGGAGGGCCATGGCATGAACCCAAAGTCTGCAGGCCATGATGCTGATGGGC 1192
QY 421 ThrGluProGlyGlnValSerSerArgSerSerSerProSerVal 435
Db 1193 TTT-----GGGGATATCTTCTCAATGAAACAAAGCAGGAGCTGTG 1231

```

RESULT 5

US-09-757-982-6
 ; Sequence 6, Application US/09757982
 ; Patent No. US20020094559A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Acton, Susan
 ; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
 ; FILE REFERENCE: WNI-050
 ; CURRENT APPLICATION NUMBER: US/09/757,982
 ; CURRENT FILING DATE: 2001-01-10
 ; PRIOR APPLICATION NUMBER: 09/163,115
 ; PRIOR FILING DATE: 1998-09-29
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 6
 ; LENGTH: 1365
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(1365)
 US-09-757-982-6

Alignment Scores:
 Pred. No.: 9.32e-33 Length: 1365
 Score: 481.00 Matches: 132
 Percent Similarity: 49.08% Conservative: 81
 Best Local Similarity: 30.41% Mismatches: 151
 Query Match: 15.96% Indels: 70
 DB: 10 Gaps: 18

US-09-830-144-2 (1-579) x US-09-757-982-6 (1-1365)

```

QY 27 AsnPhedGluIleAspTyrLysGluIleGluValGluValValGlyArgGlyAla 46
Db 19 TCCTTTGTGCAAAATAAATTTGATGACTTGCAGTTTTTTGAAAACTGCGGTGGAGGAAGT 78
QY 47 PheGlyValValCysLysAlaLysTrp-----ArgAlaLysAspValAlaIleLys--- 63
Db 79 TTTGGAGTGTATTTCGAGCCCAATGATATCACAGGACAAGGAGGTGGTGTAAAGAG 138
QY 64 -----GlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGlnLeu 81
Db 139 CTCCTCAAAATAGAGAAAGAGGCAGAA-----ATATCT 171

```

QY 82 SerArgValAsnHisProAsnIleValLysLeuTyrglyValacySLeuAsnPro----- 99
 DB 172 AGTGTCTCAGTACAGAAACATCATCTTATGAGTAATTTGAACTCCCAAC 231
 QY 100 ValCysLeuValMetGluTyrglyValagluGlySerLeuTyrglyValLeuHisgly--- 118
 DB 232 TATGGCATTTCCACAGAAATTCCTTCTTGGGATCTCATGATTAACATTAACAGTAAC 291
 QY 119 ---AlaGluProLeuProTyrglyrThrAlaAlaHisAlaMetSerTrpCysLeuGlnCys 137
 DB 292 AGAAGTACAGAGATG-----GATATGATCATCATTAATGACCTGGGCACTGATGTA 342
 QY 138 SerGlnGlyValAlaTyrglyLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeu 157
 DB 343 GCCAAAGGAATGCAATTAATTAATGATGAGAGCTCTGTCAAGGTGATTCACAGAGCCCTC 402
 QY 158 LysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGly 177
 DB 403 AAGTCAAGAAACGTTGTTATAGCTGCTGATGA---GTACTGAAGATCTGTGACTTTGGT 459
 QY 178 ThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys-----GlySerAlaAlaTrp 195
 DB 460 ---GCCCTCGGTTCATACCATTAACAACACATGCTCTGGTGGAACTTTCCCATGG 516
 QY 196 MetAlaProGluValPheGluGlySerAsnTyrglyLysCysAspValPheSerTrp 215
 DB 517 ATGGCTCCAGAAAGTTATCCAGAGTCTCCCTGTGTCAAACTGTGACACATATTCCTAT 576
 QY 216 GlyIleIleLeuTyrglyValIleThrArgArgLysProPheAspGluIleGlyGlyPro 235
 DB 577 GGTGTGTTCTTGGGAGATGCTTAACAAGGAGAGCTCCCTTTAAAGGTTTGGAGA--- 633
 QY 236 AlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgProProLeuIleLysAsn 254
 DB 634 ---TTACAATAGCTTGCTGTAGTGAAGAAAACAGAGAGATTAAACCATTCACACAGCT 690
 QY 255 LeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArg 274
 DB 691 TGCCCCAGAAAGTTTGTGAACTGTTATACATCAGTGTGGAGACCTGATGCCAAGAAACCG 750
 QY 275 ProSerMetGluGluIleValLysIleMetThrHisLeuMetArgArgTrpPheProGlyAla 294
 DB 751 CCATCATTTCAAGCAAACTATTCAATCTG----- 780
 QY 295 AspGluProLeuGlnTyrglyProCysGlnTyrglySerAspGluGlyGlnSerAsnSerAlaThr 314
 DB 781 -----GAGTCCATGTCAAAATGACAG--- 801
 QY 315 SerThrGlySerPheMetAspIleAlaSerThrAsnThrSerAsnLysSerAspThrAsn 334
 DB 802 -----AGCCTTCTTCAAGTGAAGTAACTCTTACACAAACAGGCGGAGTGAGAG 852
 QY 335 MetGluGlnValProAlaThrAsnAspThrIleLysArgLeuGlnSerLysLeuLys 354
 DB 853 TGGCAA---ATTAGAGAACTCTTGAGAGGCTTAAAGAACTAGAGCGTGAATCTCACGCTTT 909
 QY 355 AsnGlnAlaLysGlnSerGlnSerGlyArgLeuSerLeu-----GlyAlaSer 371
 DB 910 AAGGACGAGAGACTTAAAGAACGAAAGACCTTTAAAGTGTGGAGCAAAAGCTGAGA 969
 QY 372 HisGlySerSerValGlnSerLeuProProThrSer----- 383
 DB 970 GAGAGTCCAAACACCCGCTTCTTGGCTTGTGCAAGATGTCTGAGAGTCTTAC 1029
 QY 384 ---GluGlyLysArgMetSerAlaAspMetSerGluIleGluAlaArgIleAlaAlaThr 402
 DB 1030 TTGGAATCTTAAACAGAGAGTCAACAGTCAAGATGATGATGATGATGATGATGATGATGAT 1089
 QY 403 Thr---GlyAsnGlyGlnProArgArgArgSerIleGlnAspLeuThrAlaThrGlyThr 421
 DB 1090 AGTAAAGGAGGAGGCCATGAGTGAACCCCAAGTCTGAGGCCATGATGCTGATGGGCTTT 1149

QY 422 GluProGlyGlnValSerSerArgSerSerSerProSerVal 435
 DB 1150 -----GGGATATCTTCTCATATGAACAAAGCAGAGAGCTGTG 1185
 RESULT 6
 US-10-014-882-1
 ; Sequence 1: Application US/10014882
 ; Patent No. US20020107384A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hu, Yi
 ; APPLICANT: Kiehe, James
 ; APPLICANT: Donoho, Gregory
 ; TITLE OF INVENTION: No. US20020107384A1el Human Kinase and Polynucleotides Encoding r
 ; FILE REFERENCE: LEX-0279-USA
 ; CURRENT APPLICATION NUMBER: US/10/014,882
 ; PRIOR FILING DATE: 2001-12-11
 ; PRIOR FILING DATE: 2000-12-11
 ; NUMBER OF SEQ ID NOS: 3
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 3111
 ; TYPE: DNA
 ; ORGANISM: homo sapiens
 US-10-014-882-1
 Alignment Scores:
 Pred. No.: 1,37e-30 Length: 3111
 Score: 461.50 Matches: 176
 Percent Similarity: 41.16% Conservative: 94
 Best Local Similarity: 26.83% Mismatches: 238
 Query Match: 15.31% Indels: 149
 Gaps: 25
 US-09-830-144-2 (1-579) x US-10-014-882-1 (1-3111)
 QY 22 ProSerGlnValIleAsnPheGluGluIleAspTyrglySerGluIleGluValIleGluVal 41
 DB 328 CCTCGCGGCCAGCTCCCGGTACAGCTGCGCTTCAAGCGGCTGAGCTGAAGAGACTC 387
 QY 42 ValGlyArgGlyAlaPheGlyValIleValIleCysLeuValAlaSerTrpArgAlaLysAspValAla 61
 DB 388 ATCGGCGTGGGGCTTCTGGGAGGTGATCCGCCACTGGCAGGAGGCCAGAGGTGCC 447
 QY 62 IleLysGlnIleGluSerGluSerGluArgLysAlaPheIleVal----- 76
 DB 448 GTAAAGCGCGGCCAGAGACCGGACAGAGACGCGCGGCGCTGCCAGAGGTGCGG 507
 QY 77 ---GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrglyAla 95
 DB 508 CGCAGAGCTCGCTTCTTCCATGCTCGGACCCCAACATCATCGAGCTGCGGCGGTG 567
 QY 96 CysLeuAsnPro-----ValCysLeuValMetGluTyrglyValagluGlySerLeuTyrgly 113
 DB 568 TGCCTGACGAGCGCCGACCTTGTCTGTGTGAGAGTTCGCCGCGGAGAGCTCAAC 627
 QY 114 AsnValLeuHisGlyAlaGluProLeuProTyrglyrThrAla----- 127
 DB 628 CGAGCGCTGGCGGTGCAACGCGCCCGGACCGGCGCGCGCGCGCGCGCGCGCGCG 687
 QY 128 -----AlaHisAlaMet---SerTrpCysLeuGlnCysSerGlnGlyValAla 142
 DB 688 CGCGCATCCCTCCGACAGTGTGCTGCAACTGGGCGCTGACATGATGAGTGGCGGATGCTC 747
 QY 143 TyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysProProAsnLeu 162
 DB 748 TACCTGATGAGAGGCGCTTGTGCTCCATCTGACGGGAGCTCAAGTCCAGCAACATTT 807
 QY 163 LeuLeuValAlaGly-----GlyThrValLeuLysIleCysAsp 175
 DB 808 TTGCTACTTGAAGATAGAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 867
 QY 176 PheGlyThrAlaCysAsp---IleGlnThrHisMetThrAsnAsnLysGlySerAlaAla 194

[illegible]

Db 1866 TATCTTTAATAAAAAATCAGAAAAACCATTCAGCTTGGCTTCATTGTTTGGAGCCAGCCAGG 1923
 Qy 485 -ProCysProAsnSerLysGluSerMetAlaValPheGluGlnHisCysLysMetAlaG1 504
 Db 1926 GTCTGTGAAGAGCCAAAACCTTTCCCTCGATGGATTAGAA---CACAGAAAACCAAACA 1982
 Qy 504 nGluTyrMetLysValGlnThrGluIleAlaLeuLeuLeu-----GlnArgLys 520
 Db 1983 AATAAAATTGCTAGTCAGCGCTACATTGTACTTACCTCTTGGAAAAGATGCTCAGAGAGA 2042
 Qy 520 sGlnGluLeuValAlaGluLeuAspGlnAspGluLysAspGlnGlnAsnThrSerArgLe 540
 Db 2043 GAATCTTCAGAACTGAA---AGCTGGGAGGAGCGACCTCTGCGAATGTGCCACAGT 2099
 Qy 540 uValGlnGluHisLysLysLeuLeuAspGluAsnLysSerLeuSerThrTyrTyrGlnG1 560
 Db 2100 CTCATTGAG-----ATGACTCTACGATAGTCTGAGT----- 2133
 Qy 560 nCysLysGlnLeuGluValIleArgSerGlnGlnGlnLysArg 575
 Db 2134 -----AGATCCCCCAGAGAAAAGAAA 2154
 RESULT 7
 US-10-014-882-3
 ; Sequence 3, Application US/10014882
 ; Patent No. US20020107384A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hu, Yi
 ; APPLICANT: Kieke, James
 ; APPLICANT: Donoho, Gregory
 ; TITLE OF INVENTION: NO. US20020107384A1el Human Kinase and Polynucleotid
 ; FILE REFERENCE: LEX-0279-USA
 ; CURRENT APPLICATION NUMBER: US/10/014,882
 ; CURRENT FILING DATE: 2001-12-11
 ; PRIOR APPLICATION NUMBER: US 60/254,744
 ; PRIOR FILING DATE: 2000-12-11
 ; NUMBER OF SEQ ID NOS: 3
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 3518
 ; TYPE: DNA
 ; ORGANISM: homo sapiens
 US-10-014-882-3
 Alignment Scores:
 Pred. No.: 1 62e-30 Length: 3518
 Score: 461.50 Matches: 176
 Percent Similarity: 41.16% Conservative: 94
 Best Local Similarity: 26.83% Mismatches: 238
 Query Match: 15.31% Indels: 149
 DB: 12 Gaps: 25
 US-09-830-144-2 (1-579) x US-10-014-882-3 (1-3518)
 Qy 22 ProSerGlnValLeuAsnPheGluGluIleAspTyrLysGluIleGluValGluGluVal 41
 Db 591 CCCTCGGGCCAGCTCCCGGTACACGTCCGCTTCGAGCGGTGGAGCTGAAGAGCTC 650
 Qy 42 ValGlyArgGlyAlaPheGlyValValCysLysAlaLysTrpArgAlaLysAspValAla 61
 Db 651 ATCGGGCTGGGGGCTTCGGGAGAGGTGTACCGGGCCACTGGCAGGCGCAGGAGTGGCC 710
 Qy 62 IleLysGlnIleGluSerGluSerGluArgLysAlaPheIleVal----- 76
 Db 711 GTAGAGGGCGGCGCAGGACCCGGAGCAGGACGCGCGCGGTCCGAGAGCGTCCGG 770
 Qy 77 ---GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyVala 95
 Db 771 CGCAGGCTCGGCTCTTCGCCATGTCTGGCACCACCCCAACATCATCTAGCTGCGCGGCGTG 830
 Qy 96 CysLeuAsnPro-----ValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyr 113
 Db 831 TGCTGCAGCAGCGGCACCTCTGCTGCTGTGTGGAGTTGCGCCCGCGCGGAGGCTCAAC 890

```

QY 114 AsnValLeuHisGlyAlaGluProLeuProTyrThrAla----- 127
Db 891 CGAGCGTGGCGCGTCCCAAGCCCGCCGAGCCCGCGCGCGCGCGCGCG 950
QY 128 -----AlaHisAlaMet---SerTrpCysLeuGlnCysSerGlnGlyValAla 142
Db 951 CGCCGACATCCCTCCGACAGTGTGTCAACTGGCGCTGCAGATAGCGCGGGGACGTCTC 1010
QY 143 TyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysProProAsnLeu 162
Db 1011 TACCTGCATGAGAGAGCCCTTCGTGCCCATCTCCGACCGGACCTTCAGATCCAGCAACACTT 1070
QY 163 LeuLeuValAlaGly-----GlyThrValLeuLysIleCysAsp 175
Db 1071 TTGCTACTTGAGAGATAGACATGATGACATCTGCATTAACCTTTGAAGATTACAGAT 1130
QY 176 PheGlyThrAlaCysAsp---IleGlnThrHisMetThrAsnAsnLysGlySerAlaAla 194
Db 1131 TTTGGGTTGGGAGGGAATGCGACAGACACCAAAATGAGCAGACAGGACACCTATGCTC 1190
QY 195 TrpMetAlaProGluValPheGluGlnSerAsnTrpSerGlnLysCysAspValPheSer 214
Db 1191 TGGATGGCCCCCGAAGTGAATCAAGTCTTCCTTTCTTAAGGAGGACGACATCTGGAGC 1250
QY 215 TrpGlyIleIleLeuTrpGluValIleThrArgLysProPheAspGluIleGlyGly 234
Db 1251 TATGGAGTCTGCTGTGGGAACCTGCTCACCGGAGAAAGTCCCTATCGGGGACATTGATGC 1310
QY 235 ProAlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsn 254
Db 1311 CTCGGCGTGGCTTATGCGGTAGCAGCAATTAACCTCATTG---CCCATTCATCCACC 1367
QY 255 LeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArg 274
Db 1368 TGCCTCGAGCCGTTGGCCCAAGCTCATGAAGAAGATGCTGCAACAGACCTTATTTGCT 1427
QY 275 ProSer-----MetGluGlnIleValLysIleMetThrHisLeuMetArgTyr 290
Db 1428 CCATCTGTTGCCCTTATTTCTCCAGACGTTGACTGCTATTGAAGGGGAGATGATGACTGAG 1487
QY 291 PheProGlyAla-----AspGluProLeuGlnIleTyrProCysGln 303
Db 1488 ATCCCTCAGAAATCTTTATTCATTCATGCAAGATGACTGCAAACTGAAATTGCA---CA 1544
QY 304 TyrSerAspGluGlnIleSerAsnSerAlaThrSerThrGlySerPheMetArgIleAla 323
Db 1545 ATGTTTGAATGAGTTGAGAACAAAGAAAGAGAGCTCCGATCCCGGAGAGAGAGCTGACT 1604
QY 324 SerThrAsnThrSerAsnLysSerAspThrAsnMet-----GluGlnValPro 339
Db 1605 CGGGCGGCTCTGCAGAGAAAGTCTCAGAGAGAGCTCTAAAGCGCGGTGAGCAGAGCTG 1664
QY 340 AlaThrAsn-----AspThrIleLysArgLeuGlnSerLysLeuLeuLysAsnGlnAla 357
Db 1665 GCAGAGCGCGAGATTCAGCTGCTGAGCGG---GAACCTTAACATTTGATATTCAGCTA 1721
QY 358 LysGlnGlnSerLysSerLysArgLeuSerLeuGlyAlaSerHisGlySerSerValGlu 377
Db 1722 AACCGAGAGAACCCCAAGGTAAAGAAAGAGAGAGGCAAGTTTAAGAGAGTCTTTAAAG 1781
QY 378 SerLeuProProThrSerGlnGlyLysArgMetSerAlaAspMetSerGluIleGluAla 397
Db 1782 -----CTCAAGATGAGATCGAATCGAATTCAGTTTAACT---TCAGATTTCCGCAC 1826
QY 398 ArgIleAlaIleAlaThrThrGlyAsnGlyGlnProArgArgArgSerIleGlnAsnLeuThr 417
Db 1827 AAGATTAACCGTGCAGGCGCTCTCCCACTTGACAAACGCGGAGCGCTG----- 1874
QY 418 ValThrGlyThrGluProGlyGlnValSerSerArgSerSerSerProSerValArgMet 437
Db 1875 -----AACAGCAGCAGATTCCAGTCC----- 1895

```

```

QY 438 IleThrThrSerGlyProThrSerGlnLysProThrArgSerHis-ProThrThrProAs 457
Db 1896 -----CCAGAGAGCCCCCAAAAT-GATGCCCGAGCTCCGAGCCATAGACTTGCAG 1948
QY 457 PAspSerThrAspThrAsnGly----- 464
Db 1949 TGAAGCAATTAACCTTGGGGAAAGGAACACAGCTTTTGCACAGAAAGATTGAGAGATG 2008
QY 464 ----- 464
Db 2009 AAAAAGAAATTTTAAGAAAAAGTTGTACCTGGGGACCAAAATTCATTAATGAAGAA 2068
QY 465 -----SerAspAsnSerIleProMetAlaTyr 473
Db 2069 TAGAACGATTGCAAGAAAGATAGACCTCTCTCCATGAGCAACAGTCTTGGTCAAC 2128
QY 473 IleuThrIleuAspHisGly---LeuGlnProLeuAla----- 484
Db 2129 TATCTTAATTAATAAAATCAGAAAAACCATGCTTGGCTTTCATTTGTTGGACACGACAG 2188
QY 485 -ProCysProAsnSerLysGlnSerMetAlaValPheGluGlnHisCysLysMetAlaGly 504
Db 2189 GTCCGTGAAAGAGCCAAACTTCCCTGATGATTAAGAA---CACAGAAACCAAAACA 2245
QY 504 nGluTyrMetLysValGlnThrGluIleAlaLeuLeu-----GluArgLys 520
Db 2246 AATTAATTCCTTAGTCAGGCTTACATGATTAATCTTGGGAAGATGCTTAGAGACA 2305
QY 520 sGlnGluLeuValAlaGluLeuAspGlnAspGlnLysAspGlnIleAsnThrSerArgLe 540
Db 2306 GAATCTCGCAAGAGCTGAA---AGCTGGAGAGAGGAGAGCTTCGCAATGCTGCCACAT 2362
QY 540 vAlaGlnGlnHisLysLysLeuLeuAspGluAsnLysSerLeuSerThrTyrTyrGlnGly 560
Db 2363 CTCATTTAG-----ATGACTCCTACGAATAGCTGAGT----- 2396
QY 560 nCysLysLysGlnLeuGluValIleArgSerGlnGlnLysArg 575
Db 2397 -----AGATCCCCCGAGAGAAAGAA 2417

RESULT 8
US-09-947-199-1
Sequence 1, Application US/09947199
Patent No. US20020127684A1
GENERAL INFORMATION:
APPLICANT: Raju, Jeyaseelan
TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES
FILE REFERENCE: MNI-068CP2
CURRENT APPLICATION NUMBER: US/09/947,199
PRIOR FILING DATE: 2001-09-05
PRIOR APPLICATION NUMBER: 60/111,938
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: 09/291,839
PRIOR FILING DATE: 1999-04-14
PRIOR APPLICATION NUMBER: 09/458,457
PRIOR FILING DATE: 1999-12-10
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 3025
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (48)..(2552)
US-09-947-199-1

Alignment Scores:
Pred. NO.: 6,71e-27 Length: 3025
Score: 418.50 Matches: 145
Percent Similarity: 44.97% Conservative: 92
Best Local Similarity: 27.51% Mismatches: 188

```

Query Match: 13.89% Indels: 103
DB: 10 Gaps: 20

US-09-830-144-2 (1-579) x US-09-947-199-1 (1-3025)

QY	31	IleAspTyrLysGluIleGluValGluValGlyArgGlyAlaPheGlyValVal	50
Db	1419	CTTCAGCTCTCAGAATTGAGTTCCATGAGATTATGTGGCTCAGGTTCTTTGGGAAGTA	1478
QY	51	CysLysAlaLysTrpArgAlaLysAspValAlaLysGlnIleGlu-----	66
Db	1479	TATAAAGGACGATGCAGAAATAAATAGTGCTATAAAACGTTATCGAGCCAATACCTAC	1538
QY	67	---SerGluSerGluArgLysAlaPheLeValGluLeuArgGlnLeuSerArgValAsn	85
Db	1539	TGCTCCAAGTCAGATGTGGATGATGTTTGGCGAGAGGTGCCATCTCTCGCCAGCTCAAT	1598
QY	86	HisProAsnIleValLysLeuTyrGlyAlaCysLeuAsn-----ProValCysLeu	102
Db	1599	CATCCCTCGTAATTCAGTTGTGGGTGCTTGTGTAATGATCCOAGCCAGTTTGGCCATT	1658
QY	103	ValMetGluTyrAlaGluGlySerLeuTyrAsnValLeuHisGlyAlaGluProLeu	122
Db	1659	GTCACTCAATACATACAGGGGTCTCTGTCTCCCTCTCATGA-GCAGAAGAGGAT	1717
QY	123	ProTyrTyrThrAlaAlaHisAlaMetsSerTrpCysLeuGlnCysSerGln-GlyValAl	142
Db	1718	TCTTGATT-----GCAGTCTAAATTAATTATGTCAGTAGATGTGCCAAAGCATGGA	1771
QY	142	aTyrLeuHisSerMet---GlnProLysAlaLeuIleHisArgAspLeuLysPropRoAs	161
Db	1772	GTACCTTCACAACCTGACACAGCCA-----ATTATCATCGTGAACATGCATCAAA	1825
QY	161	nLeuLeuLeuValAlaGlyGlyThrValLeuLysLysCysAspPheGlyThrAlaCysAs	181
Db	1826	TATCTCTCTATGAGGATGGCATGCTGTG--GTGCCAGATTTTGAGNAATCAAAGATT	1882
QY	181	pIleGlnThr-----HisMetThrAsnAsnLysGlySerAlaAlaTrpMetal	197
Db	1883	TCTACAGTCTCTGGATGAAGAACACATGACAAAAACAACCTGGGAACCTCCGTTGGATGGC	1942
QY	197	aprogLuValPhe---GluGlySerAsnTyrSerGluLysCysAspValPheSerTrpGl	216
Db	1943	TCCTGAGGTGTTACGACGTGCATCGGTACCATCAACAGCATGCTTTCACGTATGC	2002
QY	216	yIleileLeuTrpGluValIleThrArgLysPropPheAspGluIleGlyGlyProAl	236
Db	2003	TCTGTGCTGTGGGAAATTCCTACCTGGGAAATTCCTATCGCTCATCTCAAGCCAGCGGC	2062
QY	236	aPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuleLysAsnLeuPr	256
Db	2063	TSGCGCAGCACATGGCTTACCACCAC--ATCAGACCTCCCATTGGCTATTCATTCC	2119
QY	256	oLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgProSe	276
Db	2120	CAAGCCCATATCATCTCGCTGATACGAGGTGGAACGCATGTCTCCTGAAGGAAGCCGA	2179
QY	276	rMetGluGluIleValLysIleMetThrHisLeuMetMetArgTyrPheProGlyAlaAspGl	296
Db	2180	ATTTTCTGAAGTTGTCATGAATTAGAAGATGTC-----TCCAACAT	2224
QY	296	uProLeuGlnTyrProCysGlnTyrSerAspGluGlnSerAsnSerAlaThrSerTh	316
Db	2225	TGAGCTGATGTCCTCTGCA-----TCAAGTAACAGCAG	2257
QY	316	xGlySerPheMetAspIleAlaSerThrAsn---ThrSerAsnLysSerAspThrAsnMe	335
Db	2258	TGGGTCTCTCACCTTCTCTTCTCTGATGCTGGTGAACCCGGGAGACCTGGCCG	2317
QY	335	tGluGlnValProAlaThrAsnAspThrIleLysArgLeuGluSerLysLeuLysAs	355
Db	2318	GAGTCATGTGCAGCATTAAGAAGTCGTTTCGAATTGGAATATGCTTAATGCAAGGTC	2377

Alignment Scores:

Pred. No.:	5,73e-27	Length:	2505
Score:	418.00	Matches:	120
Percent Similarity:	51.51%	Conservative:	68
Best Local Similarity:	32.88%	Mismatches:	138

US-09-947-199-3

RESULT 9

US-09-947-199-3

Sequence 3, Application US/09947199

Patent No. US20020127684A1

GENERAL INFORMATION:

APPLICANT: Raju Jeyaseelan

TITLE OF INVENTION: NOVEL DARK PROTEIN AND NUCLEIC ACID MOLECULES AND USES

TITLE OF INVENTION: THEREFOR

FILE REFERENCE: MNI-068CP2

CURRENT APPLICATION NUMBER: US/09/947,199

CURRENT FILING DATE: 2001-09-05

PRIOR APPLICATION NUMBER: 60/111,938

PRIOR FILING DATE: 1998-12-11

PRIOR APPLICATION NUMBER: 09/291,839

PRIOR FILING DATE: 1999-04-14

PRIOR APPLICATION NUMBER: 09/458,457

PRIOR FILING DATE: 1999-12-10

NUMBER OF SEQ ID NOS: 9

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 3

LENGTH: 2505

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (1)..(2505)

US-09-947-199-3


```

QY 151 AlaLeuIleHisArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrVal 170
   |||
   |||
Db 1810 ---ATCATTAACCGCGACCTCGAAGACGACCAATATTCTGCTATAGAGATGGCCATGCT 1866
QY 171 LeuLysIleCysAspPheGlyThrAlaCysAspIleGlnThr-----HisMet 186
   |||
   |||
Db 1867 GTG---GTGGCAGATTCTTGGAGAAATCAAGATTCTGAGTCCCTGGATGAGAACCAACATG 1923
QY 187 ThrAsnAsnLysGlySerAlaAlaThrMetAlaProGluValPhe---GluGlySerAsn 205
   |||
   |||
Db 1924 ACAAGACGACGAGGAACTGCGCTGGATGAGCCCTGAGGTTCACACAGTGCACGAGA 1983
QY 206 TyrSerGluLysCysAspValPheSerTrpGlyIleIleLeuTrpGluValIleThrArg 225
   |||
   |||
Db 1984 TACACCATCAAGGTGATGCTTCACTTACTCCCTGCTGAGGAGCTCCCTCACTGGA 2043
QY 226 ArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMetTrpAlaValHisAsn 245
   |||
   |||
Db 2044 GAAATTCATTGCTCATCTCAAGCCGCTGCAGACGACAGATATAGCGGTATCACACC 2103
QY 246 GlyThrArgProProLeuIleLysAsnLeuProLysProIleGluSerLeuMetThrArg 265
   |||
   |||
Db 2104 ---ATCAGACCGCGCATCGCTATTCATCCCAAGCCCAATCTCATCTGCTGATACGG 2160
QY 266 CysTrpSerLysAspProSerGlnArgProSerMetGluGluIleValLysIleMetThr 285
   |||
   |||
Db 2161 GGCCTGGAATGATGATCTCTGAAGACGACGACAGATCTCTGAAGCTTACCAACTGAG 2220
QY 286 HisLeuMetArgTrpPheProGlyAlaAspGluProLeuGlnTrpCysGlnTrpSer 305
   |||
   |||
Db 2221 GAGTGCCTA-----TGCAATGTGGAGCTCATGTCTCCAGCA----- 2256
QY 306 AspGluIleGlnSerAsnSerAlaThrSerThrGlySerPheMetAspIleAlaSerThr 325
   |||
   |||
Db 2257 -----TCAGTAACAGCAGTGGCTCTCTG----- 2280
QY 326 AsnThrSerAsnLysSerAspThrAsnMetGluGlnValProAlaThrAsnAspThrIle 345
   |||
   |||
Db 2281 TCACCTCTCTCTCTTCGATTCGATTCGCTGCTG----- 2310
QY 346 LysArgLeuGlnSerLysLeuLeuLysAsnGlnAlaLysGlnGlnSerGlnSerGlyArg 365
   |||
   |||
Db 2311 -----AGCCGG 2316
QY 366 LeuSerLeuGlyAlaSerHisGlySerSerValGlnSerLeuProProThrSerGlnGly 385
   |||
   |||
Db 2317 GGAAGGCTGCTGCGACGACCAAGTGGAGCTTACGAGAC----- 2355
QY 386 LysArgMetSerAlaAspMetSerGluIleGluAlaArgIleAlaIleThrThrGlyAsn 405
   |||
   |||
Db 2356 -----CGTTTGAAGTGGAGATGATGCCCTTAATGCAAGTCTTAAGCT 2397
QY 406 GlnGlnProArgArgArgSerIleGlnAspLeuThrValThrGlyThrGluProGly--- 424
   |||
   |||
Db 2398 GGGTGGTCCCAAGT-----GTTGGAACACACTCTAATCCGGGCGCTG 2439
QY 425 GlnValSerSerArgSerSerSerProSerValArgMetIleThrThrSerGlyProth 444
   |||
   |||
Db 2440 TCTTTGGAGAGATGAATAAGAGACCAAGTATCAACTGTGTGCAAAATAC-GGCTATGT 2498
QY 444 rSerGluLysProThrArgSerSerHisProTrpThrProAspSerThrAspThrArgGln 464
   |||
   |||
Db 2499 GTCTGATCCCATGACCGCAGCAGCATTCACTCCGCCCAACAC-----GACACCAACTT 2552
QY 464 ySerAspAsnSer----- 468
   |||
   |||
Db 2553 TGAGGACAGCAATGACAGCTCGGCATACACTAAGGGCGTCTCCCACTCAGAGCTGAC 2612
QY 469 -----IleProMetAlaIleThrLeuLeuAspHisGlnLeuGlnProLeuAlaIlePr 485
   |||
   |||
Db 2613 AGCAGTGAATTTTACCATGAGGAGGCTG-----CTTCCAAATTATTAAGCC 2657

```

```

QY 485 oCysPro 487
   |||
   |||
Db 2658 CTGGCCCT 2664

RESULT 11
US-09-947-199-9
; Sequence 9, Application US/09947199
; Patent No. US20020127684A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jayaseelan
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: THEREFOR
; CURRENT APPLICATION NUMBER: US/09/947,199
; PRIOR FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: 60/111,938
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/291,839
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 09/458,457
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 2505
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2505)
US-09-947-199-9

Alignment Scores:
Pred. No.: 1,276-26 Length: 2505
Score: 414.00 Matches: 120
Percent Similarity: 49.47% Conservative: 67
Best Local Similarity: 31.75% Mismatches: 143
Query Match: 13.74% Indels: 48
DB: 10 Gaps: 15

US-09-830-144-2 (1-579) x US-09-947-199-9 (1-2505)
QY 20 GluAlaProSerGlnValLeuAsnPhelGluGluIleAspTrpLysGluIleGluValGlu 39
   |||
   |||
Db 1351 GAACTAACCTCCCGCTTC-----CATCTCAACTCTCCGAATGAGATTCCAC 1398
QY 40 GluValValGlyArgGlyAlaPheGlyValValCysLysAlaLysTrpArgAlaLysAsp 59
   |||
   |||
Db 1399 GAGATTATCGCTCGGCTTCCTTTGGAAAGCTATAAAGGCGCATGCAAAATTAATA 1458
QY 60 ValAlaIleLysGlnIleGlu-----SerGlnSerGluArgLysAlaPhe 74
   |||
   |||
Db 1459 GTGGCATTAACGATACCGAGCCCAACACTTACTCTCAAGTCAAGAGCTGATGTCTT 1518
QY 75 IleValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTrpGly 94
   |||
   |||
Db 1519 TGCCGAGAGGTGTCATTCTGCGCAGCTCAACACCCCTGGGTGCTCAAGTTGTGGGT 1578
QY 95 AlaCysLeu-----AsnProValCysLeuValMetGluTrpAlaGluGlySer 111
   |||
   |||
Db 1579 GCGTGCCTGATGACCCCGCTGAGTTGGCATTTGCACATTAATTTCAGAGAGGCTCC 1638
QY 112 LeuTrpAsnValLeuHisGlyValaGluProLeuProTyTrpThrAlaAlaHisAlaMet 131
   |||
   |||
Db 1639 CTGTTCTCCCTGCTTAT-----GACACGAAGAATAATTTGACTTGCAGTCAATTA 1692
QY 132 SerTrpCysLeuGlnCysSerGlnGlyValAlaIleLysLeuHisSerMet--GlnProLys 150
   |||
   |||
Db 1693 ATCATTTGCGGTAGACGTTGCCAAGGAGCATGAGTACTGCACAGCTTGACCCAGCA--- 1749
QY 151 AlaLeuIleHisArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrVal 170
   |||
   |||
Db 1750 ---ATCATACCCGAGCCTGAAACAGCCCAATATTTGCTTATGAGAGTGGCATGCT 1806

```

US-09-830-144-2 (1-579) x US-09-938-842A-1014 (1-1662)

; sequence 1, Application US/09904389
; Patent No. US20020129404A1


```
Qy 102 -----LeuValMetGluTyrAlaGluGlySerLeuTyrAsnValLeuHisGly 118
Db 1609 AAATCTCCATAATCATGGAATATATCCAGAGGGAGTCTCTTCAAAATACTTCAAT 1668
Qy 119 Ala----GluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeuGlnCys 137
Db 1669 ACGAATCAGCAATG-----GACAAGAAACGCGTTTAAAGATGGCCCTTGATGT 1719
Qy 138 SerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuLeuHisArgAspLeu 157
Db 1720 GCTAGGGGAATGAATTAATCTTACACCGCAGAAATCCG---CCAATTTGTACATAGACTTG 1776
Qy 158 LysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGly 177
Db 1777 AAATCTTCCAATCTACTCGGTGGACAAGAACTGGAATGTC---AAGGTGGAGACTTTGGG 1833
Qy 178 -----ThrAlaCysAspIleGlnThrHisMetThrAsnLysGlySer 192
Db 1834 TTATCAAGTGGAGAACGCAACCTTCTTGAGTACT-----AAATCCGGGAAGGAAT 1887
Qy 193 AlaAlaTrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspVal 212
Db 1888 CCGCAGTGGATGGCTCGTGAAGTTCTCAGAAGTGAACCTTTCGAATGAGAAGTGTGATGTG 1947
Qy 213 PheSerTrpGlyIleLeuLeuTyrGluValIleThrArgArgLysProPheAspGluIle 232
Db 1948 TTCAGCTTTGGAGTCACTTATGGAGCTAATGACACGTGATACATGACATGGGACCGCTTG 2007
Qy 233 GlyGlyProAlaPheArgIleMetTrpAlaVal---HisAsnGlyThrArgProLeu 251
Db 2008 AAC-----TCTATTCAAGTTGTGGAGTTGTTGGTTTCATGATCGACGATTAGACTTA 2061
Qy 252 IleLysAsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspPro 271
Db 2062 CCTGAAGGATTAATCCCGGATCGCATCAATAATACAGGATTTGTGGCAAACTGATCCA 2121
Qy 272 SerClnArgProSerMetGluIleValLysIleMetThrHisLeuMetArgTyrPhe 291
Db 2122 GCAAAACGACCGTTCGAGGAATTAATTAGTCAGATGATGAGCCTGTTCCGCAAA--- 2178
Qy 292 ProGlyAla 294
Db 2179 CCAGGGTCA 2187
```

RESULT 15

```
US-09-938-842A-1073
; Sequence 1073, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1073
; LENGTH: 2892
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1073
Alignment Scores:
```

```
Pred. No.: 1.81e-24 Length: 2892
Score: 390.00 Matches: 104
Percent Similarity: 51.08% Conservative: 62
Best Local Similarity: 32.00% Mismatches: 96
Query Match: 12.94% Indels: 64
DB: 9 Gaps: 14
US-09-830-144-2 (1-579) x US-09-938-842A-1073 (1-2892)
Qy 22 ProSerGlnValLeuAsnPheGluGluIleAspTyrLysGluIleGluValGluVal 41
Db 1963 CTTCTTTGGCATCAGATTGGCTGGAGTCTCATGGAATGAATTCATATCAATCAAGAGCGT 2022
Qy 42 ValGlyArgGlyAlaPheGlyValValCysLysAlaLysTrpArgAlaLysAspValAla 61
Db 2033 GTGGGTGCTGGATCATTTGGAACTGTTTCATCGTCTGAGTGGCATGATGATGTTGCT 2082
Qy 62 IleLys-----GlnIleGluSerGluSerGluArgLysAlaPheIleValGlu 77
Db 2083 GTCAAGATTTGTCTATTTCAGATTTCCATGATGACCAATTCAGAGAAATTTCTCAGAGAG 2142
Qy 78 LeuArgGln-----LeuSerArgValAsnHisProAsnIleValLysLeu 92
Db 2143 GTATGTAAGCAAGCGGTTCCTATAATGAACGCTGCTCGTCAACCAATGTTGTCT-CTT 2201
Qy 93 TyrGlyAlaCys-----LeuAsnProVal-CysLeuValMetGluTyrAlaGluGly 110
Db 2202 CATGGGTCTGTGACAGAGCGACCCCGGTTATCAATAATAACAGAAATATTGTCACAGAGG 2261
Qy 110 ySerLeuTyrAsnValLeuHisGlyAlaGlu----- 120
Db 2262 CAGTCTTTTCCCTTATCCATAGCCAGCTTCTGGGGAGTTGCTAGATCAGAGGAGAG 2321
Qy 121 -----ProLeuProTyrTyrThrAlaAlaHisAlaMe 131
Db 2322 GCTACGTATGGCATTTGGATGTTGTTGCTATTCCCATTTAT----- 2364
Qy 131 tSerTrpCysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAl 151
Db 2365 -----GCCAAGGGGCTCACTACTACACTGCTCTTAATCTCT-CC 2402
Qy 151 aLeuIleHisArgAspLeuLysProAsnLeuLeuValAlaGlyGlyThrValle 171
Db 2403 TGTAGTGAATTTGGGACCTGAAATCTCCAAATCTACTGTTGATGAAGAACTGGACAGTG-- 2460
Qy 171 uLysIleCysAspPheGly---ThrAlaCysAspIleGlnThrHisMetThrAsnLys 190
Db 2461 -AAGGTTTGGCATTTTGACCTTTCAAGATTTCAAGGCAACACACTTTCATACCATCAAAATC 2519
Qy 190 s-----GlySerAlaAlaTrpMetAlaProGluValPheGluGlySerAsnTyrSer 208
Db 2520 TGTTCAGGAACACCTGAGTGGATGGCTCCAGAGTTTCTTAGAGGGGAAACCGACAAACGA 2579
Qy 208 uLysCysAspValPheSerTrpGlyIleIleLeuTrpGluValIleThrArgArgLysPr 228
Db 2580 GAATCAGATGTTTACAGTTTCGGAGTAGTCTTATGGAGTTGATTAATTTGCAACAGCC 2639
Qy 228 oPheAspGluIleGlyGlyProAlaPheArgIleMetTrpAlaValHisAsnGlyThrAr 248
Db 2640 TTGGAATGGACTC---AGTCTCTGCT---CAGGTGGTTGGAGCAGTTGTCATTCAGAAACAG 2693
Qy 248 g-----ProProLeuIleLysAsnLeuProLysProIleGluSerLeuMetTh 264
Db 2694 CGGCTTATAATTTCTTCCC-----AACACCTCTCCGCTTTTGTGATCTCTTAATGGA 2744
Qy 264 rArgCysTrpSerLysAspProSerGlnArgProSerMetGluGluIleValLysIleMe 284
Db 2745 AGCTTGTGGGCAGATGAGCGCTCTCAGCGCCAGCAATTTGGTAGTAGTAGTGGACACATT 2804
Qy 284 tThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyrProCysGlnTy 304
Db 2805 GAAGAAGCTACTAAAG-----TCTCCGGTGCAGCTGATCCAAATGGG 2846
```


Qy 304 rSerApGInGly 308
Db 2847 TGGAGCAAGGG 2859

Search completed: December 10, 2002, 02:09:05
Job time : 111.402 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Comugen Ltd.

OM protein - protein search, using sw model

Run on: December 10, 2002, 02:14:40 ; Search time 87 Seconds
(without alignments)
464.080 Million cell updates/sec

Title: US-09-830-144-2_COPY_1_303
Perfect score: 1615
Sequence: 1 MSTASASSSSSSSAGEWIE.....MTHLMRPPADEPLQPCQ 303

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1615	100.0	567	20	AA28998 Human TGF-beta act
2	1615	100.0	579	18	AA27093 Human transforming
3	1615	100.0	579	20	AA28996 Human TGF-beta act
4	1615	100.0	579	20	AA209542 Human TAK1 protein
5	1615	100.0	579	21	AA291000 Human TAK-1 protei
6	1615	100.0	579	23	AB85033 Path regulated pro
7	1615	100.0	590	20	AA209547 Human TAK1-6xHis p
8	1615	100.0	606	20	AA28997 Human TGF-beta act
9	1609	99.6	579	18	AA27092 Mouse transforming
10	813	50.3	678	22	AB58061 Drosophila melanog

11	505	31.3	252	22	AB80985 Drosophila melanog
12	434.5	26.9	367	21	AA32053 Arabidopsis thalia
13	434.5	26.9	369	21	AA32052 Arabidopsis thalia
14	434.5	26.9	407	21	AA32051 Arabidopsis thalia
15	434	26.9	349	22	AA375571 Human colon cancer
16	434	26.9	369	21	AA32172 Arabidopsis thalia
17	434	26.9	374	21	AA32171 Arabidopsis thalia
18	434	26.9	412	21	AA32170 Arabidopsis thalia
19	434	26.9	473	22	AA25382 Human protein sequ
20	433	26.8	455	21	AA318657 A human regulator
21	433	26.8	455	21	AA318657 Human survival reg
22	433	26.8	455	21	AA318657 A human regulator
23	433	26.8	800	22	AA318657 Human TGF-beta rec
24	433	26.8	800	22	AA318657 Human kinase (PKIN
25	433	26.8	1046	22	AA318657 Human kinase (PKIN
26	432.5	26.8	1021	23	ABP61000 Novel human protei
27	432	26.7	719	22	ABP61000 Human protein kina
28	432	26.7	1036	23	ABP61000 Novel human protei
29	431	26.7	1097	23	AA32177 Human PKIN-12 prot
30	427.5	26.5	341	21	AA325600 Arabidopsis thalia
31	427.5	26.5	391	21	AA325599 Arabidopsis thalia
32	427.5	26.5	1020	22	ABP61000 Drosophila melanog
33	419	25.9	847	23	AA32763 Human mitogen acti
34	411.5	25.5	589	21	AA345984 Arabidopsis thalia
35	411.5	25.5	732	21	AA345983 Arabidopsis thalia
36	411.5	25.5	760	21	AA345982 Arabidopsis thalia
37	410	25.4	835	21	AA345981 Human CARD (Cardia
38	410	25.4	835	21	AA345980 Rat CARD (Cardiac
39	410	25.4	835	22	AA345979 Novel protein kina
40	410	25.4	928	22	AB316533 Novel human diapo
41	404.5	25.0	319	21	AA325601 Arabidopsis thalia
42	401	24.8	859	16	AA325601 Human leucine-zipp
43	401	24.8	859	16	AA325601 Human leucine-zipp
44	400	24.8	888	23	AB57049 Mouse ischaemic co
45	399	24.7	977	22	AB316534 Drosophila melanog

ALIGNMENTS

RESULT 1
AA28998 standard; Protein; 567 AA.
ID AA28998 standard; Protein; 567 AA.
XX
AC AA28998;
XX
DT 29-OCT-1999 (first entry)
XX
DE Human TGF-beta activated kinase (TAK) 1c amino acid sequence.
XX
KW Nuclear factor kappa B; NF-kB; inhibitor; TGF-beta activated kinase 1;
KW TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;
KW inducible disease; atrophic dermatitis; psoriasis; viral infection;
KW endotoxin shock; septicemia; human; hTAK1c.
XX
OS Homo sapiens.
XX
PN WO9940202-A1.
XX
PD 12-AUG-1999.
XX
PF 02-FEB-1999; 99WO-JP00422.
XX
PR 30-OCT-1998; 98JP-0309316.
XX
PR 06-FEB-1998; 98JP-0026003.
XX
PA (TANA) TANA SEIYAKU CO.
XX
PI Hasegawa K, Kageyama N, Sakurai H, Sugita T;
XX
XX WPI; 1999-494298/41.
XX
XX N-PSDB; AAX99698.

PT Nuclear factor kappa B activation inhibitors, useful as preventives
 PT for, e.g. autoimmune diseases
 XX
 PS Examples; Page 43-46; 49pp; Japanese.
 XX
 CC The invention provides a method for identifying or screening a nuclear
 CC factor kappa B (NF-kB) activation inhibitor by examining the effect of a
 CC test substance on modulating the function(s) of TGF-beta activated kinase
 CC 1 (TAK1). The NFkB activation inhibitors targeting on TAK1 can be used to
 CC treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis),
 CC intractable diseases with inflammation (such as atrophic dermatitis and
 CC psoriasis), viral infection, endotoxin shock, septicemia and others. The
 CC present sequence represents the amino acid sequence of human TAK1c
 CC (hTAK1c) protein.
 XX
 XX Sequence 567 AA;
 SQ

Query Match 100.0%; Score 1615; DB 20; Length 567;
 Best Local Similarity 100.0%; Pred. No. 2e-170;
 Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTASASSSSSSAGEMIEAPSOVLNFEIDYKEIEVEEVVGRGAFGVVCKAKWRAKDV 60
 |||||
 Db 1 MSTASASSSSSSAGEMIEAPSOVLNFEIDYKEIEVEEVVGRGAFGVVCKAKWRAKDV 60
 |||||

QY 61 AIKQIESESEKAFIVELRQLSRVNHPIVKLYGACLPVCLVMEYAEAGGSLYNVLHGAE 120
 |||||
 Db 61 AIKQIESESEKAFIVELRQLSRVNHPIVKLYGACLPVCLVMEYAEAGGSLYNVLHGAE 120
 |||||

QY 121 PLPYTTAAAHAMSWCLOCSQGVAYLHSMQPKALIHRLDKPPNLLLVAGGTVLKICDFGTAC 180
 |||||
 Db 121 PLPYTTAAAHAMSWCLOCSQGVAYLHSMQPKALIHRLDKPPNLLLVAGGTVLKICDFGTAC 180
 |||||

QY 181 DIQHTMTNNKGSAAWMAPEVFEKSNYSEKCDVFSWGIIILWEVITRRKPPDEIGGPAFRIM 240
 |||||
 Db 181 DIQHTMTNNKGSAAWMAPEVFEKSNYSEKCDVFSWGIIILWEVITRRKPPDEIGGPAFRIM 240
 |||||

QY 241 WAVHNGTRPPLIKNLKPPIESLMTRCWSKDPSPQPSMEEIVKIMTHLMRYFFGADEPLQY 300
 |||||
 Db 241 WAVHNGTRPPLIKNLKPPIESLMTRCWSKDPSPQPSMEEIVKIMTHLMRYFFGADEPLQY 300
 |||||

QY 301 PCQ 303
 |||||
 Db 301 PCQ 303

RESULT 2
 AAW27093
 ID AAW27093 standard; Protein; 579 AA.
 XX
 AC AAW27093;
 XX
 DT 19-NOV-1997 (first entry)
 XX
 DE Human transforming growth factor-beta activated kinase TAK-1.
 XX
 KW TGF-beta; signal transduction; TGF-beta activated kinase;
 KW MAPK kinase activator; AMK-1; bone morphogenetic protein; BMP;
 KW protein kinase.
 XX
 OS Homo sapiens.
 XX
 XX JP09163990-A.
 PN
 PD 24-JUN-1997.
 XX
 XX 27-SEP-1996; 96JP-0256747.
 XX
 XX 24-JUL-1996; 96US-0685625.
 PR 29-SEP-1995; 95JP-0253549.
 XX
 XX (CHUS) CHUGAI PHARM CO LTD.
 PA (UENO/) UENO N.

XX WPI: 1997-380171/35.
 DR N-PSDB; AAT85095.
 XX
 PT DNA encoding transforming growth factor-beta-activated kinase, TAK-1
 PT - useful for studying the TGF-beta signal transduction system
 XX
 PS Claim 15; Page 13-15; 20pp; Japanese.
 XX
 CC The present sequence represents human transforming growth factor-beta
 CC (TGF-beta) activated kinase, TAK-1. The DNA is used to produce the
 CC TAK-1 protein which is involved in the TGF-beta family signal
 CC transduction system. TAK-1, also known as activator of MAPK kinase
 CC (AMK-1), is an enzyme which is activated by TGF-beta and bone
 CC morphogenetic protein (BMP) and activates MAPK kinase by
 CC phosphorylation.
 XX
 XX Sequence 579 AA;
 SQ

Query Match 100.0%; Score 1615; DB 18; Length 579;
 Best Local Similarity 100.0%; Pred. No. 2.1e-170;
 Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTASASSSSSSAGEMIEAPSOVLNFEIDYKEIEVEEVVGRGAFGVVCKAKWRAKDV 60
 |||||
 Db 1 MSTASASSSSSSAGEMIEAPSOVLNFEIDYKEIEVEEVVGRGAFGVVCKAKWRAKDV 60
 |||||

QY 61 AIKQIESESEKAFIVELRQLSRVNHPIVKLYGACLPVCLVMEYAEAGGSLYNVLHGAE 120
 |||||
 Db 61 AIKQIESESEKAFIVELRQLSRVNHPIVKLYGACLPVCLVMEYAEAGGSLYNVLHGAE 120
 |||||

QY 121 PLPYTTAAAHAMSWCLOCSQGVAYLHSMQPKALIHRLDKPPNLLLVAGGTVLKICDFGTAC 180
 |||||
 Db 121 PLPYTTAAAHAMSWCLOCSQGVAYLHSMQPKALIHRLDKPPNLLLVAGGTVLKICDFGTAC 180
 |||||

QY 181 DIQHTMTNNKGSAAWMAPEVFEKSNYSEKCDVFSWGIIILWEVITRRKPPDEIGGPAFRIM 240
 |||||
 Db 181 DIQHTMTNNKGSAAWMAPEVFEKSNYSEKCDVFSWGIIILWEVITRRKPPDEIGGPAFRIM 240
 |||||

QY 241 WAVHNGTRPPLIKNLKPPIESLMTRCWSKDPSPQPSMEEIVKIMTHLMRYFFGADEPLQY 300
 |||||
 Db 241 WAVHNGTRPPLIKNLKPPIESLMTRCWSKDPSPQPSMEEIVKIMTHLMRYFFGADEPLQY 300
 |||||

QY 301 PCQ 303
 |||||
 Db 301 PCQ 303

RESULT 3
 AAY28996
 ID AAY28996 standard; Protein; 579 AA.
 XX
 AC AAY28996;
 XX
 DT 29-OCT-1999 (first entry)
 XX
 DE Human TGF-beta activated kinase (TAK) 1a amino acid sequence.
 XX
 KW Nuclear factor kappa B; NF-kB; inhibitor; TGF-beta activated kinase 1;
 KW TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;
 KW intractable disease; atrophic dermatitis; psoriasis; viral infection;
 KW endotoxin shock; septicemia; human; hTAK1a.
 XX
 OS Homo sapiens.
 XX
 XX WO9940202-A1.
 PN
 XX 12-AUG-1999.
 PD
 XX 02-FEB-1999; 99WO-JP00422.
 PR 30-OCT-1998; 98JP-0309316.
 XX 06-FEB-1998; 98JP-0026003.
 PR

XX (TANA) TANABE SEIYAKU CO.
XX
XX Hasegawa K, Kageyama N, Sakurai H, Sugita T;
XX WPI: 1999-494298/41.
XX DR N-PSDB; AAX99696.
XX
XX Nucleic factor kappa B activation inhibitors, useful as preventives
XX for, e.g. autoimmune diseases
XX
XX Examples; Page 35-39; 49pp; Japanese.
XX
XX The invention provides a method for identifying or screening a nuclear
XX factor kappa B (NF-kB) activation inhibitor by examining the effect of a
XX test substance on modulating the function(s) of TGF-beta activated kinase
XX 1 (TAK1). The NFKB activation inhibitors targeting on TAK1 can be used to
XX treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis),
XX intractable diseases with inflammation (such as atrophic dermatitis and
XX psoriasis), viral infection, endotoxin shock, septicemia and others. The
XX present sequence represents the amino acid sequence of human TAK1a
XX (TAK1a) protein.
XX
XX Sequence 579 AA;
XX
XX
XX Query Match 100.0%; Score 1615; DB 20; Length 579;
XX Best Local Similarity 100.0%; Pred. No. 2.1e-170;
XX Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX
XX 1 MSTASASSSSSSAGEMIEAPSOVLNFEEDYKEIEVEEVVGAGFVGCXAKMRKADV 60
XX 1 MSTASASSSSSSAGEMIEAPSOVLNFEEDYKEIEVEEVVGAGFVGCXAKMRKADV 60
XX
XX 61 AIKQIESSEKRAFIYELRQLSRVNHPNIVKLYGACLPVCLVMEYAEGLVNVHLGAE 120
XX 61 AIKQIESSEKRAFIYELRQLSRVNHPNIVKLYGACLPVCLVMEYAEGLVNVHLGAE 120
XX
XX 121 PLPYTTAAHAMSWCLQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKICDFGTAC 180
XX 121 PLPYTTAAHAMSWCLQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKICDFGTAC 180
XX
XX 181 DIQHTMTNNKGSAAAMWAPVEFGSNYSEKCDVFSWGIIIMEVITRRKPFDEIGGPAFRIM 240
XX 181 DIQHTMTNNKGSAAAMWAPVEFGSNYSEKCDVFSWGIIIMEVITRRKPFDEIGGPAFRIM 240
XX
XX 241 WAVHNGTRPPLIKNLKPRIEISLMTKCSKDPSPRSMEIYKIMTHLMRYFPGADEPLY 300
XX 241 WAVHNGTRPPLIKNLKPRIEISLMTKCSKDPSPRSMEIYKIMTHLMRYFPGADEPLY 300
XX
XX 301 PCQ 303
XX 301 PCQ 303
XX
XX 301 PCQ 303
XX 301 PCQ 303
XX
XX
XX RESULT 4
XX AAY09542
XX ID AAY09542 standard; Protein; 579 AA.
XX
XX AAY09542;
XX
XX 21-JUL-1999 (first entry)
XX
XX Human TAK1 protein.
XX
XX Human; TAB1; TAK1; screening; inhibition; TGF-beta;
XX transforming growth factor beta.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 183..1922
XX FT /*tag= a
XX

PN W09921010-A1.
XX
XX 29-APR-1999.
XX
XX 22-OCT-1998; 98WO-JP04796.
XX
XX 22-OCT-1997; 97JP-0290188.
XX
XX (CHUS) CHUGAI SEIYAKU KK.
XX
XX Ohtomo T, Ono K, Tsuchiya M;
XX
XX WPI: 1999-312645/26.
XX DR N-PSDB; AAX56279.
XX
XX Screening for TGF- beta inhibitory substances, which are useful as
XX PT drugs for treatment of diseases relating to its disorder
XX
XX Claim 4; Page 155-157; 195pp; Japanese.
XX
XX
XX A method has been developed for screening for substances which inhibit
XX the binding of TAK1 polypeptide to TAB1 polypeptide. The method
XX comprises: (a) contacting the polypeptide in the presence of a sample;
XX and (b) detecting the amount of bound polypeptide, in which the sample
XX can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming
XX growth factor (TGF)-beta inhibitory substances can be used in drugs for
XX indications e.g. as TGF-beta signal transmission inhibitors or
XX activators, or extracellular matrix protein production enhancement
XX inhibitors or activators, or cell proliferation prevention inhibitors or
XX activators, or monocyte migration inhibitors or activators, or
XX physiological activity induction inhibitors or activators, or
XX immunosuppression inhibitors or activators, or amyloid beta protein
XX precipitation inhibitors or activators, and such substances can also be
XX inhibitors of the TAK1 polypeptide function, particularly kinase
XX activity. The present sequence represents human TAK1.
XX
XX Sequence 579 AA;
XX
XX
XX Query Match 100.0%; Score 1615; DB 20; Length 579;
XX Best Local Similarity 100.0%; Pred. No. 2.1e-170;
XX Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX
XX 1 MSTASASSSSSSAGEMIEAPSOVLNFEEDYKEIEVEEVVGAGFVGCXAKMRKADV 60
XX 1 MSTASASSSSSSAGEMIEAPSOVLNFEEDYKEIEVEEVVGAGFVGCXAKMRKADV 60
XX
XX 61 AIKQIESSEKRAFIYELRQLSRVNHPNIVKLYGACLPVCLVMEYAEGLVNVHLGAE 120
XX 61 AIKQIESSEKRAFIYELRQLSRVNHPNIVKLYGACLPVCLVMEYAEGLVNVHLGAE 120
XX
XX 121 PLPYTTAAHAMSWCLQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKICDFGTAC 180
XX 121 PLPYTTAAHAMSWCLQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKICDFGTAC 180
XX
XX 181 DIQHTMTNNKGSAAAMWAPVEFGSNYSEKCDVFSWGIIIMEVITRRKPFDEIGGPAFRIM 240
XX 181 DIQHTMTNNKGSAAAMWAPVEFGSNYSEKCDVFSWGIIIMEVITRRKPFDEIGGPAFRIM 240
XX
XX 241 WAVHNGTRPPLIKNLKPRIEISLMTKCSKDPSPRSMEIYKIMTHLMRYFPGADEPLY 300
XX 241 WAVHNGTRPPLIKNLKPRIEISLMTKCSKDPSPRSMEIYKIMTHLMRYFPGADEPLY 300
XX
XX 301 PCQ 303
XX 301 PCQ 303
XX
XX 301 PCQ 303
XX 301 PCQ 303
XX
XX
XX RESULT 5
XX AAY91000
XX ID AAY91000 standard; Protein; 579 AA.
XX
XX AAY91000;
XX

DT 04-SEP-2000 (first entry)
 XX Human TAK-1 protein sequence SEQ ID NO:2.
 DE
 XX Human; TAK-1; TAB-1; mitogen activated protein kinase; MAPK;
 KW screening; signal transduction; inhibition; inflammatory cytokine;
 KW IL-1; interleukin 1; TNF; tumour necrosis factor; inflammation;
 KW antiinflammatory; suppression.
 XX
 OS Homo sapiens.
 XX WO200023610-A1.
 PN
 XX 27-APR-2000.
 PD
 XX 21-OCT-1999; 99WO-JP05817.
 PF
 XX 21-OCT-1998; 98JP-0299962.
 PR
 XX (CHUS) CHUGAI SEIYAKU KK.
 XX
 XX Tsuchiya M, Ohtomo T, Sugamata Y, Matsumoto K;
 PI WPI; 2000-339707/29.
 XX N-PSDB; AAA39105.
 DR
 DR Method for screening inhibitors of TAK1 signal transduction for
 PT suppression of inflammatory cytokine production and use as
 PT antiinflammatory agents -
 XX
 XX Example 1; Page 80-84; 100pp; Japanese.
 PS
 XX The present invention describes a method for screening compounds for
 CC inhibition of inflammatory cytokine signal transduction by contacting
 CC the sample with TAK1 and its receptor TAB1 and selecting for inhibition
 CC of TAK1/TAB1 binding. Also described is a method for screening compounds
 CC for inhibition of inflammatory cytokine signal transduction in which the
 CC inhibition of TAK1 phosphorylation is selected for; and drug
 CC compositions for the treatment of inflammatory disorders containing as
 CC active component an inflammatory cytokine signal transduction inhibitor.
 CC TAK1 is an essential component of the signalling process which results
 CC in release of inflammatory cytokines such as interleukin-1 (IL-1),
 CC IL-10, tumour necrosis factor (TNF) and IL-6. The methods can be used
 CC for the selection of effective antiinflammatory agents. The present
 CC sequence represents human TAK-1, which is used in the exemplification of
 CC the present invention.
 XX
 XX Sequence 579 AA;
 SQ
 Query Match 100.0%; Score 1615; DB 21; Length 579;
 Best Local Similarity 100.0%; Pred. No. 2.1e-170;
 Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSTAASASSSSSSAGEMIEAPSOVLNFEIDYKEIEVEVWGRGAFGVVCKAKWRAKDV 60
 DB 1 MSTAASASSSSSSAGEMIEAPSOVLNFEIDYKEIEVEVWGRGAFGVVCKAKWRAKDV 60
 QY 61 AIKQIESESERKAFIVELRQLSRVNHNPVVKLYGACLNVPCLVMEYAEAGGSLYNVLHGA 120
 DB 61 AIKQIESESERKAFIVELRQLSRVNHNPVVKLYGACLNVPCLVMEYAEAGGSLYNVLHGA 120
 QY 121 PLPYTTAAHAMSWCLOCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKICDFTGAC 180
 DB 121 PLPYTTAAHAMSWCLOCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKICDFTGAC 180
 QY 181 DIOTHTNNKGSAAWMAPEVFGESNYSKCDVPSWGIILWEVITTRKPFDEICGGPAFRIM 240
 DB 181 DIOTHTNNKGSAAWMAPEVFGESNYSKCDVPSWGIILWEVITTRKPFDEICGGPAFRIM 240
 QY 241 WAVHNGTRPPLIKNLKPPIESLMTKRCWSKDPSPQPSMEEIVKIMTHLMRYFPGADEPLOY 300
 DB 241 WAVHNGTRPPLIKNLKPPIESLMTKRCWSKDPSPQPSMEEIVKIMTHLMRYFPGADEPLOY 300

QY 301 PCQ 303
 DB 301 PCQ 303
 RESULT 6
 ID ABB85033 standard; Protein; 579 AA.
 XX ABB85033;
 AC 16-MAY-2002 (first entry)
 XX
 DT Pain regulated protein sequence 28.
 XX
 DE Pain; analgesic; gene therapy; neurological disorder;
 KW neurodegenerative disease.
 KW
 XX Homo sapiens.
 OS
 XX WO200212338-A2.
 PN
 XX 14-FEB-2002.
 PD
 XX 03-AUG-2001; 2001WO-EP09011.
 PF
 XX 03-AUG-2000; 2000DE-1037759.
 PR
 XX (CHEF) GRUENENTHAL GMBH.
 XX
 XX Gillen C, Wetzels I, Wnendt S, Weihe E, Schaefer MK;
 PI WPI; 2002-257469/30.
 XX N-PSDB; ABL88437.
 DR
 XX Identifying pain-regulating compounds, useful for treating chronic pain
 PT and for diagnosis, by measuring binding of compounds to specific
 PT peptides and proteins -
 PT
 XX Claim 1; Fig 44; 213pp; German.
 PS
 XX The invention relates to identifying pain-regulating substances (A)
 CC comprises (i) incubating a test substance with a cell (or preparation
 CC from it) that has synthesised a peptide or protein (B) and (ii) measuring
 CC either binding of the test substance to (B) or some functional parameter
 CC that is altered by this binding. The method is useful for identifying
 CC pain-regulating substances (A) with analgesic activity. (A) along with
 CC nucleic acid (ABL88411-ABL88441) that interact with (A); (B); vectors containing the
 CC nucleic acid; antibodies against (B); cells that express (B) and agents
 CC that bind to (B), are all useful for treating pain, particularly chronic
 CC pain, including use in gene therapy. The same materials can also be used
 CC for diagnosis, e.g. of neurological and neurodegenerative diseases. The
 CC present sequence is that of a polypeptide of the invention.
 XX
 XX Sequence 579 AA;
 SQ
 Query Match 100.0%; Score 1615; DB 23; Length 579;
 Best Local Similarity 100.0%; Pred. No. 2.1e-170;
 Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSTAASASSSSSSAGEMIEAPSOVLNFEIDYKEIEVEVWGRGAFGVVCKAKWRAKDV 60
 DB 1 MSTAASASSSSSSAGEMIEAPSOVLNFEIDYKEIEVEVWGRGAFGVVCKAKWRAKDV 60
 QY 61 AIKQIESESERKAFIVELRQLSRVNHNPVVKLYGACLNVPCLVMEYAEAGGSLYNVLHGA 120
 DB 61 AIKQIESESERKAFIVELRQLSRVNHNPVVKLYGACLNVPCLVMEYAEAGGSLYNVLHGA 120
 QY 121 PLPYTTAAHAMSWCLOCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKICDFTGAC 180
 DB 121 PLPYTTAAHAMSWCLOCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKICDFTGAC 180

```

QY 181 DIQHTMTNNKGSAAWMAPEVFEKSNYSEKDVFSWGIILMEVITRRKPEDEIGGPAFRIM 240
DB 181 DIQHTMTNNKGSAAWMAPEVFEKSNYSEKDVFSWGIILMEVITRRKPEDEIGGPAFRIM 240
QY 241 WAWHNGTRPPLIKNLKPPIESIMTRCWSKDPSPQSPSMEEIVKIMTHLMRYPPGADEPIQY 300
DB 241 WAWHNGTRPPLIKNLKPPIESIMTRCWSKDPSPQSPSMEEIVKIMTHLMRYPPGADEPIQY 300
QY 301 PCQ 303
DB 301 PCQ 303

RESULT 7
AA09547
ID AA09547 standard; Protein; 590 AA.
AC AA09547;
XX 21-JUL-1999 (first entry)
DE Human TAK1-6xHis protein.
XX Human TAK1; screening; inhibition; TGF-beta;
XX transforming growth factor beta.
XX Homo sapiens.
XX Synthetic.
XX WO9921010-A1.
XX 29-APR-1999.
XX 22-OCT-1998; 98WO-JP04796.
XX 22-OCT-1997; 97JP-0290188.
XX (CHUS ) CHUGAI SEIYAKU KK.
XX Ohtomo T, Ono K, Tsuchiya M;
XX WPI; 1999-312645/26.
XX N-PSDB; AAX56285.
XX
PT Screening for TGF- beta inhibitory substances, which are useful as
XX drugs for treatment of diseases relating to its disorder
XX
PS Example 1; Page 171-174; 195pp; Japanese.
XX
CC A method has been developed for screening for substances which inhibit
CC the binding of TAK1 polypeptide to TAB1 polypeptide. The method
CC comprises: (a) contacting the polypeptide in the presence of a sample;
CC and (b) detecting the amount of bound polypeptide, in which the sample
CC can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming
CC growth factor (TGF)-beta inhibitory substances can be used in drugs for
CC indications e.g. as TGF-beta signal transduction inhibitors or
CC activators, or extracellular matrix protein production enhancement
CC inhibitors, or activators, or cell proliferation prevention inhibitors or
CC physiological activity induction inhibitors or activators, or
CC immunosuppression inhibitors or activators, or amyloid beta protein
CC precipitation inhibitors or activators, and such substances can also be
CC inhibitors of the TAK1 polypeptide function, particularly kinase
CC activity. The present sequence represents TAK1-6xHis from an example of
CC the present invention.
XX
SQ Sequence 590 AA;
Query Match 100.0%; Score 1615; DB 20; Length 590;
Best Local Similarity 100.0%; Pred. No. 2.2e-170;
Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSTAASASSSSSGAGMIAPQVLFEBIDYKEIEVEEVGARGVCAKAKRAKV 60

```

```

DB 1 MSTAASASSSSSGAGMIAPQVLFEBIDYKEIEVEEVGARGVCAKAKRAKV 60
QY 61 AIKOISESESRKAFIYELQLSRVNHPNIVKLYGACLVNCLVMEYABGSLYNNVLHGAE 120
DB 61 AIKOISESESRKAFIYELQLSRVNHPNIVKLYGACLVNCLVMEYABGSLYNNVLHGAE 120
QY 121 PLPYTAAAHAMSWCLQCSQGVAYLHSMOPKALIHRLKPPNLLVAGGVLCICPGTAC 180
DB 121 PLPYTAAAHAMSWCLQCSQGVAYLHSMOPKALIHRLKPPNLLVAGGVLCICPGTAC 180
QY 181 DIQHTMTNNKGSAAWMAPEVFEKSNYSEKDVFSWGIILMEVITRRKPEDEIGGPAFRIM 240
DB 181 DIQHTMTNNKGSAAWMAPEVFEKSNYSEKDVFSWGIILMEVITRRKPEDEIGGPAFRIM 240
QY 241 WAWHNGTRPPLIKNLKPPIESIMTRCWSKDPSPQSPSMEEIVKIMTHLMRYPPGADEPIQY 300
DB 241 WAWHNGTRPPLIKNLKPPIESIMTRCWSKDPSPQSPSMEEIVKIMTHLMRYPPGADEPIQY 300
QY 301 PCQ 303
DB 301 PCQ 303

RESULT 8
AA28997
ID AA28997 standard; Protein; 606 AA.
XX
AC AA28997;
XX 29-OCT-1999 (first entry)
XX
DE Human TGF-beta activated kinase (TAK) 1b amino acid sequence.
XX
KW Nuclear factor kappa B, NF-kB, inhibitor; TGF-beta activated kinase 1;
KW TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;
KW intracable disease; atrophic dermatitis; psoriasis; viral infection;
KW endotoxin shock; septicemia; human; hTAK1b.
XX
OS Homo sapiens.
XX
XX WO9940202-A1.
XX 12-AUG-1999.
XX 02-FEB-1999; 99WO-JP00422.
XX 30-OCT-1998; 98JP-0309316.
XX 06-FEB-1998; 98JP-0026003.
XX (TANA ) TANABE SEIYAKU CO.
XX Hasegawa K, Kageyama N, Sakurai H, Sugita T;
XX WPI; 1999-494298/41.
XX N-PSDB; AAX93697.
XX
PT Nuclear factor kappa B activation inhibitors, useful as preventives
XX for, e.g. autoimmune diseases
XX
PS Examples; Page 39-43; 49pp; Japanese.
XX
CC The invention provides a method for identifying or screening a nuclear
CC factor kappa B (NF-kB) activation inhibitor by examining the effect of a
CC test substance on modulating the function(s) of TGF-beta activated kinase
CC 1 (TAK1). The NFkB activation inhibitors targeting on TAK1 can be used to
CC treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis),
CC intracable diseases with inflammation (such as atrophic dermatitis and
CC psoriasis), viral infection, endotoxin shock, septicemia and others. The
CC present sequence represents the amino acid sequence of human TAK1b
CC (hTAK1b) protein.
XX
SQ Sequence 606 AA;

```

Query Match 100.0%; Score 1615; DB 20; Length 606;
Best Local Similarity 100.0%; Pred. No. 2.3e-170;
Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTASASSSSSSAGEMIEAPSOVLNPFEEIDYKEIEVEEVVGRGAFGVVCKAKWRAKDV 60
DB 1 MSTASASSSSSSAGEMIEAPSOVLNPFEEIDYKEIEVEEVVGRGAFGVVCKAKWRAKDV 60

QY 61 AIKQIESESEKAFIVELRQLSRVNHNPVIVKLYGACLNVPCLVMEYAEAGGSLYNVLHGAE 120
DB 61 AIKQIESESEKAFIVELRQLSRVNHNPVIVKLYGACLNVPCLVMEYAEAGGSLYNVLHGAE 120

QY 121 PLPYTAAHAMSWCLOCSQGVAYLHSMQPKALIHRLDKPPNLLLVAGGTVLKICDFGTAC 180
DB 121 PLPYTAAHAMSWCLOCSQGVAYLHSMQPKALIHRLDKPPNLLLVAGGTVLKICDFGTAC 180

QY 181 DIQTHMTNNKGSAAWMAPEVFEKGSYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIM 240
DB 181 DIQTHMTNNKGSAAWMAPEVFEKGSYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIM 240

QY 241 WAVHNGTRPPLIKNLKPKPIESLMTWCSDPSQSPSMEEIVKIMTHLMRYFPGADEPLOY 300
DB 241 WAVHNGTRPPLIKNLKPKPIESLMTWCSDPSQSPSMEEIVKIMTHLMRYFPGADEPLOY 300

QY 301 PCQ 303
DB 301 PCQ 303

RESULT 9
AAW27092
ID AAW27092 standard; Protein; 579 AA.

XX AAW27092;
AC AAW27092;
DT 19-NOV-1997 (first entry)

XX Mouse transforming growth factor-beta activated kinase TAK-1.
DE TGF-beta; signal transduction; TGF-beta activated kinase;
KW MAPK kinase activator; AMK-1; bone morphogenetic protein; BMP;
KW protein kinase.
XX Mus musculus.
XX JP09163990-A.
XX 24-JUN-1997.
XX 27-SEP-1996; 96JP-0256747.
XX 24-JUL-1996; 96US-0685625.
XX 29-SEP-1995; 95JP-0253549.
XX (CHUS) CHUGAI PHARM CO LTD.
PA (UENO/) UENO N.
XX WPI; 1997-380171/35.
DR N-PSDB; AAT85094.

XX DNA encoding transforming growth factor-beta-activated kinase, TAK-1
PT - useful for studying the TGF-beta signal transduction system
XX
XX Claim 14; Page 10-12; 20pp; Japanese.

XX The present sequence represents mouse transforming growth factor-beta
CC (TGF-beta) activated kinase, TAK-1. The DNA is used to produce the
CC TAK-1 protein which is involved in the TGF-beta family signal
CC transduction system. TAK-1, also known as activator of MAPK kinase
CC (AMK-1), is an enzyme which is activated by TGF-beta and bone
CC morphogenetic protein (BMP) and activates MAPK kinase by
CC phosphorylation.

XX SQ Sequence 579 AA;
Query Match 99.6%; Score 1609; DB 18; Length 579;
Best Local Similarity 99.7%; Pred. No. 9.8e-170;
Matches 302; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTASASSSSSSAGEMIEAPSOVLNPFEEIDYKEIEVEEVVGRGAFGVVCKAKWRAKDV 60
DB 1 MSTASASSSSSSAGEMIEAPSOVLNPFEEIDYKEIEVEEVVGRGAFGVVCKAKWRAKDV 60

QY 61 AIKQIESESEKAFIVELRQLSRVNHNPVIVKLYGACLNVPCLVMEYAEAGGSLYNVLHGAE 120
DB 61 AIKQIESESEKAFIVELRQLSRVNHNPVIVKLYGACLNVPCLVMEYAEAGGSLYNVLHGAE 120

QY 121 PLPYTAAHAMSWCLOCSQGVAYLHSMQPKALIHRLDKPPNLLLVAGGTVLKICDFGTAC 180
DB 121 PLPYTAAHAMSWCLOCSQGVAYLHSMQPKALIHRLDKPPNLLLVAGGTVLKICDFGTAC 180

QY 181 DIQTHMTNNKGSAAWMAPEVFEKGSYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIM 240
DB 181 DIQTHMTNNKGSAAWMAPEVFEKGSYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIM 240

QY 241 WAVHNGTRPPLIKNLKPKPIESLMTWCSDPSQSPSMEEIVKIMTHLMRYFPGADEPLOY 300
DB 241 WAVHNGTRPPLIKNLKPKPIESLMTWCSDPSQSPSMEEIVKIMTHLMRYFPGADEPLOY 300

QY 301 PCQ 303
DB 301 PCQ 303

RESULT 10
ABB58061
ID ABB58061 standard; Protein; 678 AA.

XX ABB58061;
AC ABB58061;
DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 975.
DE Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX Drosophila melanogaster.
XX WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US09231.
XX 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
XX (PEKE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX N-PSDB; ABL02164.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX Disclosure; SEQ ID NO 975; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL10511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (AB57737-AB572072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 678 AA;
Query Match 50.3%; Score 813; DB 22; Length 678;
Best Local Similarity 54.0%; Pred. No. 5.4e-81;
Matches 154; Conservative 45; Mismatches 82; Indels 4; Gaps 3;
QY 18 MIEAPSOVLNEEIDYKIEVEEVGAGFVCKAKRKADVAIKQIESSEKAFIVE 77
DB 1 MATASLDLQAAYVDFSSITLREKVGHSYGVCKAVWRDLVAVKFPFASAGKDIKE 60
QY 78 LROLSRVNHPVIVLYG--ACLNIVCLVMEYAEAGSLYVNLHGAEPLPYTTAAHMSWCL 135
DB 61 VKQLSRVGHPPNIIALHGSSYQOATYLMFEPAEGSLHNFJHG--KVPAYSLAHAMSMAR 119
QY 136 QCSGVAVLIHMOFKALIHRLDKPNNLLVAGTFLKICDFTACDIOTMTNNKSAAM 135
DB 120 QCAAGLAVLHMTKPLIHRYKPLNLLTNKGNLKI CDFGVADKSTMTNNGSAAM 179
QY 196 MAPEVFEGSNYSKCDVFSWGIIMEVITRRKPPDEIGPAFRIMAVHNGTRPPLIKNL 255
DB 180 MAPEVFEGSKYTEKCDIFSMAIVLMEVLSRKQPKGIDN-AVTTQMKIYGERPPLITTC 238
QY 256 PKPIESLMTKWSKDPGSRPMEIIVKIMTHLMRYFPQADEPLQY 300
DB 239 PKRIEDLMTACMKTVPEDRPSMOYIVGMHEIVKDYTGADKALEY 283
RESULT 11
ID ABB60985 standard; Protein; 252 AA.
XX
AC ABB60985;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 9747.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
XX
KM pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
XX Venter JC, Adams M, Li FWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX
DR N-PDB; ABL05088.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 9747; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL10511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (AB57737-AB572072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 252 AA;
Query Match 31.3%; Score 505; DB 22; Length 252;
Best Local Similarity 40.6%; Pred. No. 2.2e-47;
Matches 99; Conservative 48; Mismatches 71; Indels 26; Gaps 4;
QY 29 EELDYKEIEVEEVGAGFVCKAKRKADVAIKQIESSEKAFIVELRQLSRVNHFN 88
DB 6 EGVPEYELIQTKELIGTFYGSVYRAVMNRREITALKRIRGCGDKIEREIVQLTVASHVN 65
QY 89 IYKLYGACLNPCV--LVMEYAEAGSLYVNLHGAEPLPYTTAAHMSWCLQCSGVAVIHS 146
DB 66 IYELVGTSRHSCALLMEFVDGSLSSFH--AKSKPSYSHAHAFNMAHQ1AQGIAYLHG 124
QY 147 MQPKALIHRLDKPNNLLVAGTFLKICDFTACDIOTMTNNKSAAMPAE----- 199
DB 125 MQPKAVIHHDIPMLTLTLCEKGLKIKICDFGVVDLSOSISNAGTCYKAPVEVLELDF 184
QY 200 -----VFEGSNYSKCDVFSWGIIMEVITRRKPPDEIGPAFRIMAVH 244
DB 185 KENRIINOPTFOKVLQGNKPKDCKDVYSMAITFWEILSRKEPEQY-NTLFELYMALN 243
QY 245 NGTR 248
DB 244 EGR 247
RESULT 12
ID AAG32053 standard; Protein; 367 AA.
XX
AC AAG32053;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 38597.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX
XX termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
XX
PR 05-MAR-1999; 99US-0123180.
XX
PR 09-MAR-1999; 99US-0123548.
XX
PR 23-MAR-1999; 99US-0125788.
XX
PR 25-MAR-1999; 99US-0126284.
XX
PR 29-MAR-1999; 99US-0126785.
XX
PR 01-APR-1999; 99US-0127462.
XX
PR 06-APR-1999; 99US-0128234.
XX
PR 08-APR-1999; 99US-0128714.
XX
PR 16-APR-1999; 99US-0129845.
XX
PR 19-APR-1999; 99US-0130077.
XX
PR 21-APR-1999; 99US-0130449.
XX
PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132488.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 23-APR-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 26-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156559.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.

PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 26.9%; Score 434.5; DB 21; Length 367;

Best Local Similarity 34.8%; Pred. No. 2.8e-39; Indels 55; Gaps 13;

Matches 110; Conservative 54; Mismatches 97; Indels 55; Gaps 13;

QY 12 SSSAGEMIEAPSOY-----LNFEET-IDYKEIEVEEVRGGA 46
DB 37 SLASVGSVFRGRTHALNDALQALMDTRYPTREGLTINDEWTDIRKLNMGPAFAQGA 96
QY 47 FGVCCKAKRAKDVAIKQIE--SESEKRA-----FTVEERQTSRVNHPNIVKLYGACLN 98
DB 97 FGKLYKGTYNEDVAIKILERPENSPEKAQFMEOQFOEVSMLANLKHPRNIVPRIGACRK 156
QY 99 PV--CLWMEVYEGSGSLNVL---HGAEPLPYTAAHAMSCLOCSGVAYLHSMQKAL 152
DB 157 PMWVCIVTEYVAKGGSVROFLTRQNRVPLKL-----AVKQALDVAGMAVHG--RNF 208
QY 153 IHRDLKPNLLVAGGVYKICDPGTA-CDIQTH-MTNKGSAAWMAPEVEGSGNYSEKC 210
DB 209 IHRDLKSNLILISADKSI-KIADFGVARIEVQTEGMPETICTYMMAPBEMIQHRAVYQKV 267
QY 211 DVFSWGIILWEVITRRKPFDEIGG--PAFRIMAVHNGTRPPLIKNLKPPIESIMTRCWS 268
DB 268 DVYSGFGLWMLITGLLPFQNMVAVQAFAV---VNRGVRTVENDCLPYLSDIMTRCWD 324
QY 269 KDPSQRPSEMEIVKIM 284
DB 325 ANPEVRPCFEVVKLL 340

RESULT 13

AAG32052
ID AAG32052 standard; Protein; 369 AA.

XX AAG32052;

DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 38596.

XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

OS Arabidopsis thaliana.

XX EPI033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX

PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.

```
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 17-JUL-1999; 99US-0144086.
PR 18-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.

PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 26.9%; Score 434.5; DB 21; Length 369;
Best Local Similarity 34.8%; Pred. No. 2.9e-39;
Matches 110; Conservative 54; Mismatches 97; Indels 55; Gaps 13;

QY 12 SSSAGEMIEAPSQV-----LNFE--IDYKEIEVEEVVGRGA 46
Db 39 SLVGSQSVFRGRVTHALDDALQAALMDPRYPTTEGLTNYDEWTIDLRKLNMGPAFAQGA 98
QY 47 FGVCYKAKWAKDVAIKQIB--SESEKKA-----FIVELRQLSRVNHNPVVKLYGACLN 98
Db 99 FGKLYKGTNGEDVAIKILERPENSPEKAQFMEQOQFQEEVSMLANLKHNPVIFIGACRK 158
QY 99 PV--CLVMEYAEAGGSLYNVL---HGAEPUPYYTAHAHMSWCLQCSQGVAYLHSMQPKAL 152
Db 159 PMWCIVTEYAKGSGVRQFLTRRQNRVAVLKL-----AVKQALDVARGMAYVHG---RNF 210
QY 153 IHRDLKPPNLLVAGGTVLKICDFGTA-CDIQTH-MTNKNGSSAAMNAPEVFECSNYSEK 210
Db 211 IHRDLKSDNLLISADKSI-KIADFGVARIEVQTEGTPETGTYRWAPENIQHRAYNQKV 269
QY 211 DVPSWGIILWEVITRRKPPDEIGG--PAFRIMWAVHNGTRPPLIKNLKPPIESLMTRCWS 268
Db 270 DVYSFGIVLWELITGLLPFQNMATAVQAFAV---VNRGVRPTVPNDCLPVLSDIMTRCWD 326
QY 269 KDSQRPSEMEIVKIM 284
Db 327 ANPEVRPCFVEVVKLL 342

RESULT 14
AAG32051
ID AAG32051 standard; Protein; 407 AA.
XX AAG32051;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 38595.
```

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
OS Arabidopsis thaliana.
XX EP1033405-A2.
PN
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0125264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132485.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.


```
QY 131 MSWCLQCSQGVAYLHSMQPKALIHRLKPEPNLLVAGTVLKICDEGTACDIQTHMTNK 190
Db 164 MTWATDVAKGMHYLHMEAPVKYIHRDLKSRNVIAADG-VLKICDFG-ASRFHNHTHMS 221
QY 191 --GSAAMMAPEVEEGSNYSEKCDVFESWGIILWEVITRRKPFDEIGGPAFRIMW-AVHNGT 247
Db 222 LVGTFFPMAPEVIQSLPVSETCDTYSYGVVLMEMLTREVFPKGLEG--LQYAMLVEKNE 279
QY 248 RPPLINKLPKPIESLMTRCMSKPSQSPSMEEIVKIM 284
Db 280 RLTIPESSCPSPFAELHQCMEADAKRPSFKQIISIL 316
```

Search completed: December 10, 2002, 03:47:53
Job time : 89 secs

C;Accession: JCS955
R;Sakurai, H.; Shigemori, N.; Hasegawa, K.; Sugita, T.
Biochem. Biophys. Res. Commun. 243, 545-549, 1998
A;Title: TGF-beta-activated kinase 1 stimulates NF-kappaB activation by an NF-kappaB-ind
A;Reference number: JCS955; MUID:98153801; PMID:9480845
A;Accession: JCS955
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-579 <SAK>
A;Cross-references: DDBJ:AB009356; NID:g2924623; PIDN:BAA25025.1; PID:g2924624
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C;Keywords: phosphotransferase

Query Match 100.0%; Score 1615; DB 2; Length 579;
Best Local Similarity 100.0%; Pred. No. 1.9e-83;
Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTAASASSSSSSAGEMIEAPSOVLNFEIDYKEIEVEEVVGRGAFGVVCKAKWRAKDV 60
DB 1 MSTAASASSSSSSAGEMIEAPSOVLNFEIDYKEIEVEEVVGRGAFGVVCKAKWRAKDV 60
QY 61 AIKOIESESRKAFIVELRQLSRVNHPIVVKLYGACLNVPCLVMEYAEAGGSLYNVLHGAE 120
DB 61 AIKOIESESRKAFIVELRQLSRVNHPIVVKLYGACLNVPCLVMEYAEAGGSLYNVLHGAE 120
QY 121 PLPYTTAAHAMSWCLQCSQGVAYLHSMQPKALIHRLDKPPNLLLVAGGTVLKICDFGTAC 180
DB 121 PLPYTTAAHAMSWCLQCSQGVAYLHSMQPKALIHRLDKPPNLLLVAGGTVLKICDFGTAC 180
QY 181 DIQTHMTNNKGSAAWMAPEVFGSNYSEKCDVFSWGIIWEVITRRKPFDEIGGPAFRIM 240
DB 181 DIQTHMTNNKGSAAWMAPEVFGSNYSEKCDVFSWGIIWEVITRRKPFDEIGGPAFRIM 240
QY 241 WAVHNGTRPPLIKNPKPIESLMTRCWSKDPQSRPSMEEIVKIMTHLMRYFPGADPQY 300
DB 241 WAVHNGTRPPLIKNPKPIESLMTRCWSKDPQSRPSMEEIVKIMTHLMRYFPGADPQY 300
QY 301 PCQ 303
DB 301 PCQ 303

RESULT 3
JCS956
transforming growth factor-beta activated kinase (EC 2.7.-.-) 1b - human
C;Species: Homo sapiens (man)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C;Accession: JCS956
R;Sakurai, H.; Shigemori, N.; Hasegawa, K.; Sugita, T.
Biochem. Biophys. Res. Commun. 243, 545-549, 1998
A;Title: TGF-beta-activated kinase 1 stimulates NF-kappaB activation by an NF-kappaB-ind
A;Reference number: JCS955; MUID:98153801; PMID:9480845
A;Accession: JCS956
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-606 <SAK>
A;Cross-references: DDBJ:AB009357; NID:g2924625; PIDN:BAA25026.1; PID:g2924626
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C;Keywords: phosphotransferase

Query Match 100.0%; Score 1615; DB 2; Length 606;
Best Local Similarity 100.0%; Pred. No. 2e-83;
Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTAASASSSSSSAGEMIEAPSOVLNFEIDYKEIEVEEVVGRGAFGVVCKAKWRAKDV 60
DB 1 MSTAASASSSSSSAGEMIEAPSOVLNFEIDYKEIEVEEVVGRGAFGVVCKAKWRAKDV 60
QY 61 AIKOIESESRKAFIVELRQLSRVNHPIVVKLYGACLNVPCLVMEYAEAGGSLYNVLHGAE 120
DB 61 AIKOIESESRKAFIVELRQLSRVNHPIVVKLYGACLNVPCLVMEYAEAGGSLYNVLHGAE 120
QY 121 PLPYTTAAHAMSWCLQCSQGVAYLHSMQPKALIHRLDKPPNLLLVAGGTVLKICDFGTAC 180

DB 121 PLPYTTAAHAMSWCLQCSQGVAYLHSMQPKALIHRLDKPPNLLLVAGGTVLKICDFGTAC 180
QY 181 DIQTHMTNNKGSAAWMAPEVFGSNYSEKCDVFSWGIIWEVITRRKPFDEIGGPAFRIM 240
DB 181 DIQTHMTNNKGSAAWMAPEVFGSNYSEKCDVFSWGIIWEVITRRKPFDEIGGPAFRIM 240
QY 241 WAVHNGTRPPLIKNPKPIESLMTRCWSKDPQSRPSMEEIVKIMTHLMRYFPGADPQY 300
DB 241 WAVHNGTRPPLIKNPKPIESLMTRCWSKDPQSRPSMEEIVKIMTHLMRYFPGADPQY 300
QY 301 PCQ 303
DB 301 PCQ 303

RESULT 4
S68178
mixed-lineage protein kinase 2 (EC 2.7.1.-) - human
C;Species: Homo sapiens (man)
C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 11-Jun-1999
C;Accession: S68178; I38044; S32468
R;Dorow, D.S.; Devereux, L.; Tu, G.F.; Price, G.; Nicholl, J.K.; Sutherland, G.R.; Simpe
Eur. J. Biochem. 234, 492-500, 1995
A;Title: Complete nucleotide sequence, expression, and chromosomal localisation of humar
A;Reference number: S68178; MUID:96128179; PMID:8536694
A;Accession: S68178
A;Molecule type: mRNA
A;Residues: 1-954 <DOR>
A;Cross-references: EMBL:X90846; NID:g971419; PIDN:CAA62351.1; PID:g971420
R;Kato, M.; Hirai, M.; Sugimura, T.; Terada, M.
Oncogene 10, 1447-1451, 1995
A;Title: Cloning and characterization of MST, a novel (putative) serine/threonine kinase
A;Reference number: I38044; MUID:95249256; PMID:7731697
A;Accession: I38044
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-461, 'A', 'V', 465-470, 'S', 472-806 'R', 808-817, 'A', 819-954 <RES>
A;Cross-references: EMBL:Z48615; NID:g758592; PIDN:CAA85531.1; PID:g758593
R;Dorow, D.S.; Devereux, L.; Dietzsch, E.; de Kretser, T.
Eur. J. Biochem. 213, 701-710, 1993
A;Title: Identification of a new family of human epithelial protein kinases containing t
A;Reference number: S32467; MUID:93238756; PMID:8477742
A;Accession: S32468
A;Molecule type: mRNA
A;Residues: 244-464, 'AQAAGRRQHPALWL' <DO2>
C;Genetics:
A;Gene: GDB:MLK2; GDB:MST
A;Cross-references: GDB:362654; GDB:624810; OMIM:600137
A;Map position: 19q13.1-19q13.2
C;Superfamily: mixed-lineage protein kinase 2; protein kinase homology; SH3 homology
C;Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific protein k
F;23-76/Domain: SH3 homology <SH3>
F;96-364/Domain: protein kinase homology <KIN>
F;104-112/Region: protein kinase ATP-binding motif
F;384-405/Region: leucine zipper motif
F;419-440/Region: leucine zipper motif
F;449-463/Region: basic
F;125,145,222,224/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 29.9%; Score 482.5; DB 1; Length 954;
Best Local Similarity 38.4%; Pred. No. 5.3e-20;
Matches 107; Conservative 51; Mismatches 98; Indels 23; Gaps 7;

QY 21 APSQVLNFEIDYKEIEVEEVVGRGAFGVVCKAKWRAKDVVAIKQIESESRKAFIV---- 76
DB 83 APAGLQLPQEIPTPHLQLEIIIGVGFGKVVYRVALWRGEEVAVKAARLDPEKDPAVTAEQV 142
QY 77 --ELRQLSRVNHPIVVKLYGACLNVP--VCLVMEYAEAGGSLYNVLHGAEPLPYTTAAHAMS 132
DB 143 CQEARLFGALQHPNIIALRGACLNPPHCLVMEYAEAGGALSRLAGRRVPPHV----LVN 198
QY 133 WCLQCSQGVAYLHSMQPKALIHRLDKPPNLLLV-----AGGTVLKICDFGTACD-IQT 184

Db	199	WAVQVARGNNYLHNDA	PVPIITHRDLKSNILILEA	ENHNHLADTLKTTDGLAREMHT	258
Qy	185	HMTNNKGSAAAMAP	EEVEFGSNSEKCDVFSWII	IMEVITRRKPEDEIGCPAFRIMNAV	243
Db	259	TKMSAAGYVAMMA	PEVIRLSTLSFKSDVMSFGVLL	MELTGEVPRE--DALNAVYVA	316
Qy	244	HNGTRPPLIKNL	PKPIESLMTRCWSKDP	QSRPSMEIYX	282
Db	317	MNKLTLPI	PSCTCEPFA	RLLECWDPHGRD	FGSLIK 355

RESULT 5
G84635
Probable protein kinase [imported] - Arabidopsis thaliana
C,Species: Arabidopsis thaliana (mouse-ear cress)
C,Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 17-May-2002
A,Accession: G84635
R,Link, X.: Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vankken, S.E.; Unayam, L.; Tallon, L.;
Euss, D.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A,Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A,Reference number: A84420; MUID:20083487; PMID:10617197
A,Accession: G84635
A,Status: Preliminary
A,Molecule type: DNA
A,Residues: 1-407 <STO>
A,Cross-references: GB:AE002093; NID:g4337195; PIDD:AA01809.1; GSPDB:GN00139
C,Genetics:
A,Gene: Atg242460
A,Map position: 2
A,Superfamily: kinase-related transforming protein; protein kinase homology

Query Match	26.9%	Score 434.5	DB 2	Length 407
Best Local Similarity	34.8%	Pred. No. 1.2e-17		
Matches 110; Conservative	54	Mismatches 97	Indels 55	Gaps 13

```

Oy 12 SSSAGMEIARSQV-----LNFE--IDYKEI EVEEVRGGA 46
Db 77 SLVSQSVFRRGRVTHALNDLALQALMDTRYPTGELTNDEWTIDIRKLNMGPAFOGA 136
Oy 47 FGVCVAKMRADVAIKOIE--SESEKA-----FIVEIKLSRVNHPINIVKLYGACLN 98
Db 137 FGKLKYGTYNGEDAIKILERPENSPEKAQMEQOFOOEVSMLNLHPNIVRTIGACRK 196
Oy 99 PV--GLVMEVYBEGSLVYVL-----HGAEPILPYTTAAHMSWCLCSQGVATLHMQKAL 152
Db 197 PMWVCITVEIYAKGSSVRPFLTRQKRAVPLK-----AVKALDVAAGMAVYHG--RNF 248
Oy 153 IHRDLKPNLLLVAGTIVLKI CDPEGTA-CDIQTH-MTNNKSAAAMAPEVEGGSYSEKC 210
Db 249 IHRDLKSNLLISADKSI-KIADPEVAALIEVQTSGMTPESTYRMAPEMIQHAYNOKV 307
Oy 211 DVSNGITLMEVITRRKPEDEIG--PAFRIMAAVHNGTRPPLIKNPKPIESIMTRCWS 268
Db 308 DVSNGIVIMEHITGLTFPQNMNTAQAALFAV--VNRGVRPTVBNDCLPVLSDIMTRCWD 364
Oy 269 KDPSSQPSMEELVKIM 284
Db 365 ANPEVRPCFEVEVVKLL 380

```

RESULT 6
T10671
protein kinase homolog FE21.90 - *Arabidopsis thaliana*
C:Species: *Arabidopsis thaliana* (mouse-ear cress)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 19-May-2000
C:Accession: T10671
R:Beran, M.; Lennard, N.; Quail, M.; Harris, B.; Ralandream, M.A.; Barrell, B.G.; Bancroft
submitted to the Protein Sequence Database, June 1999
A:Reference number: Z16533
A:Accession: T10671

A: Molecule type: DNA
A: Residues: 1-412 <BEV>
A: Cross-references: EMBL:AL049911; GSPDB:GN00062; ATSP:F6E21.90
A: Experimental source: cultivar Columbia; BAC clone F6E21
C: Genetics:
A: Gene: ATSP:F6E21.90
A: Map position: 4
A: Introns: 300/2
C: Superfamily: Kinase-related transforming protein; protein kinase homology
F13-392/domain: protein kinase homology <KIN>

[illegible]

RESULT 7
T18287
protein-tyrosine kinase (EC 2.7.1.112) - slime mold (Dictyostelium discoideum)
C.Species: Dictyostelium discoideum
C.Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Feb-2000
C.Accession: T18287
R.Adler: K.
submitted to the EMBL Data Library, July 1996
A.Reference number: Z18856
A.Accession: T18287
A.Status: Preliminary; translated from GB/EMBL/DBJ
A.Molecule type: DNA
A.Residues: 1-1338 <ADL>
A.Cross-references: EMBL:U04830; NID:gl1468982; PID:gl1468983; PIDN:AA04999.1
C.Genetics:
A.Introns: 1181/3
C.Keywords: ATP; phosphotransferase; tyrosine-specific protein kinase

Query Match	26.6%	Score 430;	DB 2;	Length 1338;
Best Local Similarity	37.7%	Pred. NO. 6e-17;		
Matches 98; Conservative	50;	Mismatches 94;	Indels 18;	Gaps 7

```

QY 29 EIDYKELEVEEVGRAGVYCCAKKRRADVAIKQIESE--SERKAFIVELRQLSRVNH 86
Db 1050 KEIKDEVAIVIERVGAQSPFNVSLGIWNGYGAIAKITLKNEISINDKFLKEVSSLIKSYH 1105
QY 87 PNIYLVACINPVCILMVEYAEAGSLVNVYH----GAEPILPYTTAAHAWSWICQSQGYA 142
Db 1110 PNVVTFMGARIDPPCITFEYIQGSLVDVHTQIKIKLNPIMYKMIHDL-----SLGME 1165
QY 143 YLHMKOPKALIHRLDKPEPNLLVAGGVTLKICDFGTACDIQTHMT--NNKGSAAWMAPEYF 201
Db 1164 HHHSTQ--MHRLDTSKNILLDBFNKI-KIAPFGIATTLSDMTLSGTRNRMSPBLT 1211
QY 202 EGSNSYSEKDVFSNGIILMEVITRRKPEDEIGPAPRIMAVHNGTRPPLIKLKPYES 261
Db 1220 KGLVYNEKVADVYSGLVVYEIVTGKIPFEEISDTAAANAFAFEN-RRPAIPDPDVSJLEK 1276

```

QY 262 LMTFCWSKDPSPQSRPSMEEIV 281
Db 1279 LITKCWASDPSPQSRPSFTEIL 1298

RESULT 8
T48115
protein kinase ATMRL1 (EC 2.7.1.1-) [imported] - Arabidopsis thaliana
N;Alternate names: protein Fl6M2.110
C;Species: Arabidopsis thaliana (mouse-ear cross)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2001
C;Accession: T48115; T51942
R;Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Mewes, H.W.; Rudd, S.; Lemcke, K.
submitted to the Protein Sequence Database, April 2000
A;Reference number: Z24459
A;Accession: T48115
A;Molecule type: DNA
A;Residues: 1-391 <RIE>
A;Cross-references: EMBL:AL138648; NID:g7523397; PIDN:CAB86427.1; PID:g7523408
A;Experimental source: cultivar Columbia; BAC clone Fl6M2
R;Ichimura, K.; Mizoguchi, T.; Shinozaki, K.
Plant Sci. 130, 171-179, 1997
A;Title: ATMRL1, an Arabidopsis protein kinase related to mammal mixed-lineage kinases a
A;Reference number: Z24427
A;Accession: T51942
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-391 <ICH>
A;Cross-references: EMBL:AB006810; NID:g2351096; PIDN:BAA22079.1; PID:g2351097
A;Experimental source: cultivar Columbia
C;Genetics:
A;Map position: 3
A;Introns: 109/1; 149/3; 220/2; 278/3; 323/3
A;Note: Fl6M2.110
C;Superfamily: Kinase-related transforming protein; protein kinase homology
C;Keywords: phosphotransferase

Query Match 26.5%; Score 427.5; DB 2; Length 391;
Best Local Similarity 31.3%; Pred. No. 2.9e-17;
Matches 103; Conservative 58; Mismatches 101; Indels 67; Gaps 10;

QY 2 STASASSSSSAGEMIEAPSOVLNPEEIDYKEIEVEEVGGRGAFGVCKAKWRAKDYA 61
Db 55 STSVLNTSASTS-----APAMQEWELDLKMDKHVLHAGTGTGTVRGVYAGQEVA 108

QY 62 IKQI-----ESESERKAFIVELRQLSRVNHNPVVKLYGACL----- 97
Db 109 VKVLDMGEDGYATPAETALRASFEQEVAVWQKLDHPNVTKFIGASWGTSDLRIPAGDT 168

QY 98 -----NPVCLVMEYAEGLSLNVLHGAEPLPYTA-----AHAMSWCLQCSQGVAY 143
Db 169 GGRGNGAHPARACCVVEYVAGGTLKKFL-----IKKYRAKLPIKDVIQLALDLARGLSY 223

QY 144 LHSMQPKALIHRLKPNLLLVAGGTVLKICDFGTA---CDIQTHMTNNKSGAAWMAPEV 200
Db 224 LHS---KAIVHRDVKENMMLLPQNKTLKIADFGVARVAQNPQDMGTGTGLGYWAPEV 279

QY 201 FEGSNYSEKCDVFSWGIILWEVITRRKP-----FDEIGGPAFRIMWVHNGTRPPLIKNL 255
Db 280 LEGKPNRKCDVYSFGVCLWEIYCCDMPYADCSEFABISHA-----VVHRNLRPEIPKCC 333

QY 256 PKFIESLMTFCWSKDPSPQSRPSMEEIVKIM 284
Db 334 PHAVANIMKRCWDPNDRRPEMEVVKLL 362

RESULT 9
C84856
probable protein kinase [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cross)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: C84856

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: C84856
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-357 <STO>
A;Cross-references: GB:AE002093; NID:g4559329; PIDN:AAD22991.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g42630
A;Map position: 2

Query Match 26.1%; Score 421.5; DB 2; Length 357;
Best Local Similarity 36.3%; Pred. No. 5.9e-17;
Matches 106; Conservative 50; Mismatches 85; Indels 51; Gaps 15;

QY 17 EMIEAP----SOVLNPEE--IDYKEIEVEEVGGRGAFGVCKAKWRAKDYAIK-----QI 65
Db 79 EVLESMPQNKPLLPFFBEWNIDFSKLKVGASVSGGTSGVVGCRGVWYKTEVAIKIFLQQQL 138

QY 66 ESESERKAFIVELRQLSRVNHNPVVKLYGACLN-----VCLVMEYAEGLSLNVLHGAEPLP 123
Db 139 TAEN-MKVFCNEISILSRLOHPNVLILGACTKPPQLSLVTEYMWSTGSLDVR----- 191

QY 124 YTTAAHMSW-----CLQCSQGVAYLHMQPKALIHRLKPNLLLVAGGTVLKICDFG 177
Db 192 --TRKELLSWQKLILAEICRGLMYIHKM--GIVHRDLTSAACLL--NKSIVKICDFG 244

QY 178 TACDIQTHMTNN-----KGSAAWMAPEVEGSGNYSYSEKCDVFSWGIILWEVITRRKPPD 230
Db 245 ----LSRRMTGTAVKDTAAAGTPEWMAPELRNPEVPTKSDIFSFGVIMMELSTLSKPMK 300

QY 231 BIGGPAFRIMWVHN-GTRPPLIKNLPKPIESLMTRCWSKDPSPQSRPSMEEIV 281
Db 301 --GVPEKVIHIVANEGAR---LKIPEGPLQKLIADCS-EPEQRPSCKEIL 346

RESULT 10
A53800
mixed-lineage protein kinase (EC 2.7.1.1-) 3 - human
N;Alternate names: protein kinase PTK1; protein kinase SPRK
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A53800; I58395
R;Gallo, K.A.; Mark, M.R.; Scadden, D.T.; Wang, Z.; Gu, Q.; Godowski, P.J.
J. Biol. Chem. 269, 15092-15100, 1994
A;Title: Identification and characterization of SPRK, a novel src-homology 3 domain-cont
A;Reference number: A53800; MUID:94253068; PMID:8195146
A;Accession: A53800
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-847 <GAL>
A;Cross-references: GB:U07747; NID:g464027; PIDN:AAA19647.1; PID:g464028
R;Ring, Y.L.; Leung, I.W.; Heng, H.H.; Tsui, L.C.; Lassar, N.J.
Oncogene 9, 1745-1750, 1994
A;Title: MLK-3: identification of a widely-expressed protein kinase bearing an SH3 domai
A;Reference number: I58395; MUID:94239754; PMID:8183572
A;Accession: I58395
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-847 <RES>
A;Cross-references: GB:L32976; NID:g488295; PIDN:AAA59859.1; PID:g488296
C;Genetics:
A;Gene: GDB:MLK3; PTK1; SPRK
A;Cross-references: GDB:134755; OMIM:600050
A;Map position: liq3.1-1lq3.3
C;Superfamily: mixed-lineage protein kinase 3; protein kinase homology; SH3 homology
C;Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific protein k
F;48-100/Domain: SH3 homology <SH32>
F;115-383/Domain: protein kinase homology <KIN>

F:123-131/Region: protein kinase ATP-binding motif
F:403-424/Region: leucine zipper motif
F:438-459/Region: leucine zipper motif
F:468-482/Region: basic

Query Match 25.9%; Score 419; DB 1; Length 847;
Best Local Similarity 34.6%; Pred. No. 1.7e-16;
Matches 102; Conservative 53; Mismatches 102; Indels 38; Gaps 9;

Qy 22 PSCVNFEEIDYKEIEVEEVGAGFVCAKWRKADVAIK-----QISESESR 71
Db 108 PCVNASFC-----ELRLFEVIGIGFGKXKRGSRKGLVAVKAAODPPEDISVTAESR 162
Qy 72 KAVIVLRQLSRVNHNYKLYGACINP--VCIWVEYAGSSLVNVLHGAEPLPYTAAH 129
Db 163 Q-----EARLFAMLAHPNIIALKAVCLEBEPNLCVMEYAGGPLSRALAGRVRPHV---- 214
Qy 130 AMSWCICQSGVAYLHSMQPKALIHRLDKPNNLLV-----AGTVLKICDFTGACD 181
Db 215 LVNNAVQIRGMHYLHCEALVPVIRHDKSNLILLQPIESDMEHKTKITDFTGLAREW 274
Qy 182 IQTHMTNKGSAAMAPEVEGSGNSYSEKCDVFSWGIIMEVITRRKPFDEIGGPAFRIMW 241
Db 275 HKTTQMSAAGTAVAMMAPEVITKSTSKSDVMSFGVLMELLTGEVPIRGIDCLAVANGV 334
Qy 242 AVHNGTRPPLIKNLKPIESLMTRCMSKDPSCQPS---MEIIVKIMTHMKRYFP 292
Db 335 AVNKLTL-PIPTCEPFAQLMADCAODPHRRDPFASITLQLEALAEQVLRMP 368

RESULT 11

JU0229
mixed-lineage protein kinase 1 - human
C/Species: Homo sapiens (man)
C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 17-Mar-1999
C/Accession: S32467; JU0229
R/Dow, D.S.; Devereux, L.; Dietzsch, E.; de Kretser, T.
Eur. J. Biochem. 213, 701-710, 1993
A>Title: Identification of a new family of human epithelial protein kinases containing
A/Reference number: S32467; MUID:93238756; PMID:8477742
A/Accession: S32467
A/Molecule type: mRNA
A/Residues: 1-394 <DO2>
C/Genetics:
A:Gene: GDB:MLK1
A/Cross-references: GDB:141921; OMIM:600136
A/Map position: 14q24.3-14q31
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C/Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific protein k
F:1-269/Domain: protein kinase homology <KIN>
F:1-268/Domain: catalytic <CAT>
F:9-17/Region: protein kinase ATP-binding motif
F:289-310/Region: leucine zipper motif
F:324-345/Region: leucine zipper motif
F:354-368/Region: basic

Query Match 25.9%; Score 418.5; DB 2; Length 394;
Best Local Similarity 36.6%; Pred. No. 9.4e-17;
Matches 98; Conservative 47; Mismatches 100; Indels 23; Gaps 8;

Qy 35 EIEVEEVGAGFVCAKWRKADVAIK--QISESESRKAFIYELNQ-----LSRVNHPN 88
Db 2 ELTLEIIGIGFGKXKRGSRKGLVAVKAAODPPEDISVTAESR 162
Qy 89 IVKLYGACINP--VCIWVEYAGSSLVNVLHGAEPLPYTAAHMSQCICQSGVAYLHS 146
Db 62 IIAIRGVCLKEPNLCVMEFARGPLNRVLSGKRIPDI-----LVNNAVQIARGMVLIHD 117
Qy 147 MQPKALIHRLDKPNNLLV-----AGTVLKICDFTGACD-IQTHMTNKGSAAMAP 198
Db 118 EAIYPIIHRDKSNLILLQVENGDISNKLKITDFTGLAREWRTTKMSAAGTAVAMAP 177
Qy 199 EVFEGSNYSEKCDVFSWGIIMEVITRRKPFDEIGGPAFRIMAV--HNGTRPPLIKNLK 257

Db 178 EVIAMSFSKSGDVMSYGVLLMELLTGEVFRGIDG--LRAVAGVAMNKLALPISTCPE 235
Qy 258 PIESLMTRCMSKDPSCQPSMEIIVKIMT 285
Db 236 PFAKLMECWNPDPHSRPSFTNILDQLT 263

RESULT 12

h16747
hypothetical protein R13F6.7 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Mar-2000
C/Accession: T16747
R/Miller, N.
submitted to the EMBL Data Library, April 1994
A/Description: The sequence of C. elegans cosmid R13F6.
A/Reference number: Z18570
A/Accession: T16747
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1328 <ML>
A/Cross-references: EMBL:U00046; NID:9470358; PID:9470364; PIDN:AA047047.1; GSPDB:GN0002
A/Experimental source: strain Bristol N2; clone R13F6
C/Genetics:
A:Gene: CESP:R13F6.7
A/Map position: 3
A/Introns: 20/3; 160/3; 222/2; 286/2
C/Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3 h

Query Match 25.8%; Score 416; DB 2; Length 328;
Best Local Similarity 33.4%; Pred. No. 1.1e-16;
Matches 99; Conservative 57; Mismatches 112; Indels 28; Gaps 11;

Qy 1 MSTRASASSSSSAGBMTAPSOVNFEEIDYKEIEVEEVGAGFVCAKWRKADV 60
Db 22 MSTPTSNESSTSSSNN-----DQVLPFDIQRDDIQVDHIGVGFAGVFGSNWTLPDG 76
Qy 61 AIHQISESESRKAFIYELNQ--LSRVNHPNIVKLYGACN--NPVCIWVEYAGSSLVNV 115
Db 77 SQRTI--ALKKVFLEKEAEILSKIRHKNIIQFYGICATGNDPFIYVEAKESLYDF 133
Qy 116 LHGAEPPLPYTAAHMS-----WCLQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGT 169
Db 134 IHSEESQSFASSSGNSPDVYVWMAQIASGIQLHYDAVDITIIHRDLKSKVVL-DKVL 192
Qy 170 VLKICPFTACDIQTHMTNKK--GSAAMAPD-VFEGSNYSEKCDVFSWGIIMEVITR 225
Db 193 VCKICDFGTSKDL-THSCAPSWGTAAMMSPEMLIQSEGLTTATDVMSYGVLLMEILSK 251
Qy 226 RKPFDEIGGPAFRIMAV--HNGTRPPLIKNLKPIESLMTRCMSKDPSCQPSMEI 280
Db 252 EVYKYXS--EFRITMTQSGITTLAIPSPCAPLKQMSNCWMTPKDRANMROI 305

RESULT 13

D84555
probable protein kinase (imported) - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C/Accession: D84555
A:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Berrilo, M.I.; Town, C.D.; Fujii, C.Y.; N
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vankken, S.E.; Umayan, L.; Tallon, L.
eaus, D.; Niernan, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: A84420; MUID:20083487; PMID:10617197
A/Accession: D84555
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-546 <STO>
A/Cross-references: GB:AEO02093; NID:96598802; PIDN:AA080785.2; GSPDB:GN00139
C/Genetics:

A:Gene: At2g17700
A:Map position: 2

Query Match 25.5%; Score 411.5; DB 2; Length 546;
Best Local Similarity 33.3%; Pred. No. 3e-16;
Matches 96; Conservative 61; Mismatches 116; Indels 15; Gaps 8;

QY 10 SSSSAGEMIEAPSOVLNFEIDYKEIEVEVVGAFGVCKAKWRAKDVAIKQIES 69
DB 260 SNEILPACIEIPTDGTDEWEIDVTQLKIEKKVAGSYGDLHRTGYCQEVAKFLKPD 319
QY 70 ERKAFIVELRQ-----LSRVNHNPNIVKLYGACLN--PVCVMEYAEAGSGLYNVLHCAEPLP 123
DB 320 VNNEMLREFSQEVFMKRVKRVNIVQFLGACTRSPCLTCLVTFMARGSIYDFLHKQKCA- 378
QY 124 YTTAAHANSWCLQCSGVAYLHSMQPKALIHRLDKPPLLVLVAGGTVLKICDFGTA-CDI 182
DB 379 -FKLQTLTKVALDVAKGMSYLHQ-----NNIIHRDLKTANLLMDEHGLV-KVADFGVARVQI 433
QY 183 QTH-MTNNKGSAAWMAPEVFEGSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMW 241
DB 434 ESGVMTAETGTYRWMAPEVIEHKPYNHRKADVFSYAIVLWELLTGDIPIYAFV-TPLQAAVG 492
QY 242 AVHNGTRPPLIKNLPKPIESLMTRCWSKDPQSPQSMEEIVKIMTHLMR 289
DB 493 VVQGLRPKIPKTKTHPKVKGLLERCHWDPEQRPUFEEIEMLQOIMK 540

RESULT 14

T04683
hypothetical protein F8D20.290 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 21-May-1999
C:Accession: T04683
R:Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Jesse, T.; Heijnen, L.; Vos, P.; Mewes, submitted to the Protein Sequence Database, July 1998
A:Reference number: Z15381
A:Accession: T04683
A:Molecule type: DNA
A:Residues: 1-553 <BEV>
A:Cross-references: EMBL:AL031135
A:Experimental source: cultivar Columbia; BAC clone F8D20
C:Genetics:
A:Map position: 4
A:Introns: 69/2; 107/3; 176/2; 194/3; 231/3; 289/2; 325/2; 350/1; 408/3; 440/3; 467/3; 4
A:Note: F8D20.290

Query Match 25.5%; Score 411.5; DB 2; Length 553;
Best Local Similarity 32.2%; Pred. No. 3.1e-16;
Matches 96; Conservative 60; Mismatches 111; Indels 31; Gaps 9;

QY 19 IEAPSOVLNFEIDYKEIEVEVVGAFGVCKAKWRAKDVAIKQIESERKAFIVEL 78
DB 258 VEIPTDGTDEWEIDMKQLKIEKKVAGSYGELFRGTGYCQEVAKILKPERVNAEMLREF 317
QY 79 RQ-----LSRVNHNPNIVKLYGACLN--PVCVMEYAEAGSGLYNVLHCAEPLPYTTAAHMS 132
DB 318 SQEYIMRKVRHKRVNIVQFAGCTRSPNLCIVTFMTRGSIYDFLHKHKGV--FKIOSLLK 375
QY 133 WCLQCSGVAYLHSMQPKALIHRLDKPPLLVLVAGGTVLKICDFGTACDIQTH---MTNN 189
DB 376 VALDVSKGMNYLHQ-----NNIIHRDLKTANLLM-DEHVVVKVADFGVA-RVQTESGVMTAE 430
QY 190 KGSAAWMAPEVFEGSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMWVHNGTRP 249
DB 431 TGTYRWMAPEVIEHKPYDHRADVFSYAIVLWELLTGELPYSYL-TPLQAAVGVVQKGLRP 489
QY 250 PLIKNLPKPIESLMTRCWSKDPQSPQSMEEIVKIMTHLMR-----YFPG 293
DB 490 KIPKETHPKLTLELCKWCQDPAALRPFAETIEMLNQLIREVIDLSLHKDKHGGYFSG 547

F96763
hypothetical protein F25P22.8 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: F96763
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.P.; Hughes, B.; Huijar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ket, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: F96763
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1030 <STO>
A:Cross-references: GB:AE005173; NID:G6692730; PIDN:AAF24836.1; GSPDB:GN00141
C:Genetics:
A:Gene: F25P22.8
A:Map position: 1

Query Match 25.4%; Score 411; DB 2; Length 1030;
Best Local Similarity 32.9%; Pred. No. 5.5e-16;
Matches 103; Conservative 57; Mismatches 125; Indels 28; Gaps 10;

QY 1 MSTASAASSSSSSSAGEMIEAPSOVLNFEIDYKEIEVEVVGAFGVCKAKWRAKDV 60
DB 721 ISDKSIGNESSKSDCDDVSDC-----ELLWEITVGERIGLSYGEVYRGDMHGTEV 772
QY 61 AIKQIESE----SERKAFIVELQRLSRVHNPNIVKLYGACLN--PVCVMEYAEAGSGLYN 114
DB 773 AVKFKLDQLTGEALBEFRSEVRIMKLRHPNIVLFMGAVTRPPNLSIVTEFLPGRSLYR 832
QY 115 VLHGAELPLPYTTAAHMSWCLQCSGVAYLHSMQPKALIHRLDKPPLLVLVAGGTVLKIC 174
DB 833 LIH--RPNQLDERRRLMALDAARGNMYLHSCNP-MIVHRDLKSPN-LLVDRKNWVVKVC 888
QY 175 DFGTACDIQTHMT-----NNKGSAAWMAPEVFEGSNYSEKCDVFSWGIILWEVITRRKPF 229
DB 889 DFGLS--RMKHSTVLSKSTAGTAGTAENWMAPEVLRNEPADEKCDVSYGVILWELFTLQQPW 946
QY 230 DEIGGPAFRIMWVHNGTRPPLIKNLPKP-LESIMTECWSKDPQSPQSMEEIVKIMTHLM 288
DB 947 GKMN--PMQVVGVAVGFOHRRLDIPDFVDPALADLSKWCQTDLSKLRPSFAEIMASLRLQ 1004
QY 289 RYFFGADEPLQYP 301
DB 1005 KPTVGSNIPRPVP 1017

Search completed: December 10, 2002, 03:51:45
Job time : 55 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 10, 2002, 03:06:45 ; Search time 43 Seconds

(without alignments)
292.263 Million cell updates/sec

Title: US-09-830-144-2_COPY_1_303

Perfect score: 1615

Sequence: 1 MSTASASSSSSSSAGEMIR.....MTHLMRYPGADEPLQYPCQ 303

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

112892

Total number of hits satisfying chosen parameters:

112892

Minimum DB seq length: 0

2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length DB	ID	Description
1	1615	100.0	606	1	M3K7_HUMAN
2	1609	99.6	579	1	M3K7_MOUSE
3	486	30.1	393	1	M3K7_DROME
4	482.5	29.9	954	1	M3KA_HUMAN
5	418.5	25.9	394	1	M3K9_HUMAN
6	408.5	25.3	410	1	KYK2_DICDI
7	403	25.0	888	1	M3KC_RAT
8	401	24.8	859	1	M3KC_HUMAN
9	398	24.6	888	1	M3KC_MOUSE
10	376.5	23.3	821	1	CTK1_ARATH
11	365.5	22.6	1584	1	KYK1_DICDI
12	365.5	22.6	630	1	TEC_MOUSE
13	358.5	22.2	631	1	TEC_HUMAN
14	356	22.0	620	1	ITTK_HUMAN
15	354.5	22.0	625	1	ITTK_MOUSE
16	354.5	22.0	527	1	TKX_MOUSE
17	354	21.9	984	1	EPA3_RAT
18	352	21.8	1114	1	RET_HUMAN
19	349	21.6	983	1	EPA3_HUMAN
20	348.5	21.5	1115	1	RET_MOUSE
21	346.5	21.5	804	1	FPS_DROME
22	344.5	21.3	983	1	EPA3_CHICK
23	343	21.2	467	1	MATK_RAT
24	341.5	21.1	590	1	BTXL_DROME
25	341.5	21.1	981	1	EPA3_BRARE
26	338.5	21.0	981	1	EPA3_MOUSE
27	337	20.9	983	1	EPA3_MOUSE
28	336.5	20.8	983	1	EPA7_CHICK
29	335	20.7	980	1	MML_AV1MH
30	334.5	20.7	980	1	EPA7_RAT
31	333.5	20.7	444	1	KI28_CAEEL
32	333.5	20.7	604	1	KRAA_MOUSE
33	333.5	20.7	604	1	KRAA_RAT

34	333.5	20.7	606	1	KRAA_PIG	Q19004 sus scrofa
35	333.5	20.7	998	1	EPA7_HUMAN	Q15375 homo sapien
36	333	20.6	507	1	MATK_HUMAN	P42679 homo sapien
37	332	20.6	367	1	MML_AV1I1	P10533 avian sapien
38	332	20.6	450	1	RML1_AVEVR	P27966 avian rous-
39	332	20.6	765	1	KRAH_HUMAN	P15056 homo sapien
40	332	20.6	806	1	RML1_CHICK	Q04982 gallus gall
41	332	20.6	807	1	RML1_COTUA	P34908 coturnix co
42	331	20.5	402	1	KROS_AV1SU	P00529 avian sarco
43	331	20.5	648	1	KRAF_RAT	P11345 rattus norv
44	331	20.5	822	1	FRG1_RAT	Q04589 rattus norv
45	330.5	20.5	328	1	KRAH_MOUSE	P28028 mus musculu

ALIGNMENTS

```

RESULT 1
M3K7_HUMAN          STANDARD;          PRT;          606 AA.
ID   O43318; O43317; O43319;
AC   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Mitogen-activated protein kinase kinase kinase 7 (EC 2.7.1.-)
DE   (Transforming growth factor-beta-activated kinase 1) (TGF-beta-
DE   activated kinase 1).
GN   MAP3K7 OR TAK1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX   NCBI_Taxid=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC   TISSUE=Lung;
RA   MEDLINE=98153801; PubMed=9480845;
RX   Sakurai H., Shigemori N., Hasegawa K., Sugita T.;
RT   "TGF-beta-activated kinase 1 stimulates NF-kappa B activation by an
RT   NF-kappa B-inducing kinase-independent mechanism.";
RL   Biochem. Biophys. Res. Commun. 243:545-549(1998).
CC   FUNCTION: CAN PHOSPHORYLATE AND ACTIVATE YBT UNDEFINED MAPKKS.
CC   -!- MEDIATOR OF TGF-BETA SIGNAL TRANSDUCTION. STIMULATES NF-KAPPA B
CC   ACTIVATION.
CC   -!- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1A, 1B (SHOWN HERE) AND 1C; ARE
CC   PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC   MAP KINASE KINASE KINASE SUBFAMILY.
CC   -!- SIMILARITY: IN THE C-TERMINAL SECTION; STRONG, TO C210RF7.
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (see http://www.ebi.ac.uk/announcements
CC   or send an email to license@ebi.ac.uk).
EMBL; AB009357; BAA25026.1; -
EMBL; AB009356; BAA25025.1; -
EMBL; AB009358; BAA25027.2; -
HSSP; P08631; 1AD5.
GeneW; HGNC:6859; MAP3K7.
MIM; 602614; -
InterPro; IPR000719; Euk_kinase.
InterPro; IPR004040; Ser_kinase.
InterPro; IPR002290; Ser_thr_kinase.
Pfam; PF00069; Kinase; 1.
ProDom; PD000001; Euk_kinase; 1.
SMART; SMO0221; STYK; 1.
PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
Transferase; Serine/threonine-protein kinase; ATP-binding;

```

```
KW Alternative splicing.
FT DOMAIN 8 14
FT NP_BIND 36 291
FT BINDING 42 50
FT ACT_SITE 63 63
FT VARSPLIC 156 156
FT VARSPLIC 404 430
FT VARSPLIC 509 518
FT VARSPLIC 519 606
SQ SEQUENCE 606 AA; 67196 MW; 308F8147CD174013 CRC64;

Query Match 100.0%; Score 1615; DB 1; Length 606;
Best Local Similarity 100.0%; Pred. No. 7e-134;
Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTAASASSSSSSAGEMIEAPSOVLNFEEDYKEIEVEEVVGRGAFGVVCKAKWRAKDV 60
DB 1 MSTAASASSSSSSAGEMIEAPSOVLNFEEDYKEIEVEEVVGRGAFGVVCKAKWRAKDV 60
QY 61 AIKQIESESEKAFIVELRQLSRVNHPIVLYGACLNPCVCLVMEYAEAGGSLYNNVLHGAE 120
DB 61 AIKQIESESEKAFIVELRQLSRVNHPIVLYGACLNPCVCLVMEYAEAGGSLYNNVLHGAE 120
QY 121 PLPYTTAAHAMSWCLQCSQGVAYLHSMQKALIHRLDKPNNLLVAGGTVLKCDFGTAC 180
DB 121 PLPYTTAAHAMSWCLQCSQGVAYLHSMQKALIHRLDKPNNLLVAGGTVLKCDFGTAC 180
QY 181 DIQTHMTNNKGAAMWAPVFEKSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIM 240
DB 181 DIQTHMTNNKGAAMWAPVFEKSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIM 240
QY 241 WAVHNGTRPPLIKNLKPKIESLMTKCSKDPQSPMSMEEIVKIMTHLMRYFFPGADEPLQY 300
DB 241 WAVHNGTRPPLIKNLKPKIESLMTKCSKDPQSPMSMEEIVKIMTHLMRYFFPGADEPLQY 300
QY 301 PCQ 303
DB 301 PCQ 303

RESULT 2
M3K7_MOUSE STANDARD; PRT; 579 AA.
ID M3K7_MOUSE
AC 062073;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mitogen-activated protein kinase kinase kinase 7 (BC 2.7.1.-)
DE (transforming growth factor-beta-activated kinase 1) (TGF-beta-
DE activated kinase 1).
GN MAP3K7 OR TAK1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96123277; PubMed=8533096;
RA Yamaguchi K., Shirakabe K., Shibuya H., Irie K., Ohishi I., Ueno N.,
RA Taniguchi T., Nishida E., Matsumoto K.;
RT Identification of a member of the MAPKK family as a potential
RT mediator of TGF-beta signal transduction."
RL Science 270:2008-2011(1995).
CC -!- FUNCTION: CAN PHOSPHORYLATE AND ACTIVATE YET UNDEFINED MAPKKS.
CC -!- MEDIATOR OF TGF-BETA SIGNAL TRANSDUCTION. STIMULATES NF-KAPPA B
CC ACTIVATION.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -!- MAP KINASE KINASE SUBFAMILY.
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; STRONG, TO C21ORF7.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```

```
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D76446; BAA11184.1; -.
CC HSSP; P08631; LADS.
CC MGD; MGI:1346877; Map3k7.
CC InterPro; IPR000719; Euk_pkinase.
CC InterPro; IPR004040; STY_pkinase.
CC InterPro; IPR002290; Ser_thr_pkinase.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Euk_pkinase; 1.
CC SMART; SM00221; STYKC; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
CC TRANSFERASE; Serine/threonine-protein kinase; ATP-binding.
FT DOMAIN 8 16
FT NP_BIND 36 291
FT BINDING 42 50
FT ACT_SITE 63 63
FT ACT_SITE 156 156
SQ SEQUENCE 579 AA; 64227 MW; 97C8F6F3C8E283EE CRC64;

Query Match 99.6%; Score 1609; DB 1; Length 579;
Best Local Similarity 99.7%; Pred. No. 2.2e-133;
Matches 302; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTAASASSSSSSAGEMIEAPSOVLNFEEDYKEIEVEEVVGRGAFGVVCKAKWRAKDV 60
DB 1 MSTAASASSSSSSAGEMIEAPSOVLNFEEDYKEIEVEEVVGRGAFGVVCKAKWRAKDV 60
QY 61 AIKQIESESEKAFIVELRQLSRVNHPIVLYGACLNPCVCLVMEYAEAGGSLYNNVLHGAE 120
DB 61 AIKQIESESEKAFIVELRQLSRVNHPIVLYGACLNPCVCLVMEYAEAGGSLYNNVLHGAE 120
QY 121 PLPYTTAAHAMSWCLQCSQGVAYLHSMQKALIHRLDKPNNLLVAGGTVLKCDFGTAC 180
DB 121 PLPYTTAAHAMSWCLQCSQGVAYLHSMQKALIHRLDKPNNLLVAGGTVLKCDFGTAC 180
QY 181 DIQTHMTNNKGAAMWAPVFEKSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIM 240
DB 181 DIQTHMTNNKGAAMWAPVFEKSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIM 240
QY 241 WAVHNGTRPPLIKNLKPKIESLMTKCSKDPQSPMSMEEIVKIMTHLMRYFFPGADEPLQY 300
DB 241 WAVHNGTRPPLIKNLKPKIESLMTKCSKDPQSPMSMEEIVKIMTHLMRYFFPGADEPLQY 300
QY 301 PCQ 303
DB 301 PCQ 303

RESULT 3
M3K7_MOUSE STANDARD; PRT; 393 AA.
ID M3K7_MOUSE
AC P83104;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative mitogen-activated protein kinase kinase 7
DE (EC 2.7.1.-).
GN TAK1L1
OS Drosophila melanogaster (Fruit fly)
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
```


RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amatiades P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortan J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Chame M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abrial J.F., Agayanti A., An H.-J., Andrews-Plamkovich C., Baldwin D.,
 RA Ballew R.M., Basu A., Bakendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
 RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahike C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart M.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kemielsen J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Mostrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venner E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP CONCEPTUAL TRANSLATION.
 RA Manning G., Sudarshan S., Pionman G.;
 RT "Prediction of novel protein kinases from the *Drosophila* genome
 project and EST sequences.";
 RL Unpublished observations (Aug-2001).
 CC -!- FUNCTION: Can phosphorylate and activate yet undefined MAPKs.
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC MAP KINASE KINASE KINASE SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AE003732; -!- NOT ANNOTATED_CDS.
 DR FlyBase: FBgn0046689; Tak1.
 DR InterPro: IPR000719; Euk_Pkinase.
 DR InterPro: IPR002290; Ser_thr_kinase.
 DR InterPro: IPR001245; Tyr_kinase.
 DR Pfam: PF00069; Pkinase.1.
 DR PRINTS: PR00109; TYRKINASE.
 DR ProDom: PD000001; Euk_Pkinase.1.
 DR SMART: SM00220; S_TKc; 1.
 DR SMART: SM00219; TyKc; 1.
 DR PROSITE: PS00107; PROTEIN KINASE ATP.1.
 DR PROSITE: PS00101; PROTEIN KINASE DOM.1.
 DR PROSITE: PS00108; PROTEIN KINASE ST.1.
 KW Hypothetical protein: Transferase; Serine/threonine-protein kinase;
 KW Tyrosine-protein kinase; ATP-binding.
 FT DOMAIN 11 266 PROTEIN KINASE.
 FT NP_BIND 17 25 ATP (BY SIMILARITY).

FT BINDING 38 38 ATP (BY SIMILARITY).
 FT ACT_SITE 133 133 BY SIMILARITY.
 SQ SEQUENCE 393 AA; 45238 MW; 43370C98490F0384 CRC64;
 Query Match 30.1%; Score 486; DB 1; Length 393;
 Best Local Similarity 36.5%; Pred. No. 3.3e-35;
 Matches 105; Conservative 66; Mismatches 91; Indels 24; Gaps 9;
 QY 29 EIDYKELEV-EEVVGKAPGVCKAKRMDAIK-----QIESESEKATIVELRQ 80
 DB 3 KQVDAEYKLESEKFLGAGSGGAVRKATFQNEIAVKIFDFLEETIKKNAER-----EITH 57
 QY 81 LSRVNHPIVLKYGACLN--PVLVMEYAEAGSGSYNVLHGAEPPLPYTAAHMSWCQCS 138
 DB 58 LSEIDHEVIVIGVIGASGKKDYLLMELEBGSJAHNYLGYDKNK--YVEQAVRWALQCA 116
 QY 139 QGVAVLHSMOPKALIHRLKRPNLVLVAGGVVLICPGFACDIDTHTNKGSAAMAP 198
 DB 117 KALAVLHSLD-RPIVHRDIKQNMMLVQHEDLKICDFGLATDSNNKTDMDGLRTMAP 175
 QY 199 EVFEGSNYSKCDVFSWGIILMEVITRRKPDDEIGP--AFRIMAVHNGTRPPL---IK 253
 DB 176 EAIKHLKKTATACDVSFQIMMELMTRLQPSHLNPNQVAYIKALSSGKELPMNAVRS 235
 QY 254 NLPKPIESLMTRCNSKSDSPSMEIYKINTHLMRYFPGADEPLQIP 301
 DB 236 DPEBGIKQIMECCMDINPEKRPMSKEIEKFLGE--QVESGTDEDFIXP 281
 RESULT 4
 ID K3XA HUMAN STANDARD; PRT; 954 AA.
 AC Q02779; Q12761; Q14871;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Mitogen-activated protein kinase kinase kinase 10 (EC 2.7.1.37)
 DE (Mixed lineage kinase 2) (protein kinase MST).
 GN MAP3K10 OR MLK2 OR MST.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 ON NCBI_taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=96128179; PubMed=8536694;
 RA Dorow D.S., Devereux L., Tu G.F., Price G., Nicholl J.K.,
 RA Sutherland G.R., Simpson R.J.;
 RT "Complete nucleotide sequence, expression, and chromosomal
 RL localisation of human mixed-lineage kinase 2.";
 RL Eur. J. Biochem. 234:492-500(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=95249256; PubMed=7731697;
 RA Katoh M., Hirai M., Sugimura T., Terada M.;
 RT "Cloning and characterization of MST, a novel (putative)
 RL serine/threonine kinase with SH3 domain.";
 RL Oncogene 10:1447-1451(1995).
 RN [3]
 RP SEQUENCE OF 244-480 FROM N.A.
 RC TISSUE=colon epithelium;
 RX MEDLINE=93238756; PubMed=8477742;
 RA Dorow D.S., Devereux L., Dietzsch E., de Kreijter T.;
 RT "Identification of a new family of human epithelial protein kinases
 RT containing two leucine/isoleucine-zipper domains.";
 RL Eur. J. Biochem. 213:701-710(1993).
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN BRAIN AND SKELETAL MUSCLE.
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC MAP KINASE KINASE KINASE SUBFAMILY.
 CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.

```
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X90846; CAA62351.1; -.
CC EMBL; Z48615; CAA88531.1; -.
CC PIR; S32468; S32468.
CC HGSP; P11362; IFGK.
CC Genew; HGNC:6849; MAP3K10.
CC MTM; 600137; -.
CC InterPro; IPR000719; Euk_pkinase.
CC InterPro; IPR001452; SH3.
CC InterPro; IPR004040; STY_pkinase.
CC InterPro; IPR002290; Ser_thr_pkinase.
CC InterPro; IPR001245; Tyr_pkinase.
CC Pfam; PF00018; SH3_1.
CC Pfam; PF00069; pkinase; 1.
CC PRINTS; PR00452; SH3DOMAIN.
CC PRINTS; PR00109; TYRKINASE.
CC ProDom; PD000001; Euk_pkinase; 1.
CC ProDom; PD000066; SH3_1.
CC SMART; SM00326; SH3_1.
CC SMART; SM00221; SYTK; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS50002; SH3_1.
CC Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase;
KW ATP-binding; SH3 domain.
KW
FT DOMAIN 2 5 POLY-GLU.
FT DOMAIN 16 81 SH3.
FT DOMAIN 98 360 PROTEIN_KINASE.
FT NP_BIND 104 112 ATP (BY SIMILARITY).
FT BINDING 125 125 ATP (BY SIMILARITY).
FT ACT_SITE 222 222 BY SIMILARITY.
FT DOMAIN 384 405 LEUCINE-ZIPPER 1 (BY SIMILARITY).
FT DOMAIN 419 440 LEUCINE-ZIPPER 2 (BY SIMILARITY).
FT DOMAIN 449 463 ARG/LYS-RICH (BASIC).
FT CONFLICT 462 464 SRL -> AV (IN REF. 2).
FT CONFLICT 465 480 LKREGGSHSLPSGF -> AQAGRRQHPQALWL (IN
FT REF. 3).
FT CONFLICT 471 471 G -> S (IN REF. 2).
FT CONFLICT 807 807 G -> R (IN REF. 2).
FT CONFLICT 818 818 V -> A (IN REF. 2).
FT SEQUENCE 954 AA; 103623 MW; 538F4AAA559B0ABA CRC64;
Query Match 29.9%; Score 482.5; DB 1; Length 954;
Best Local Similarity 38.4%; Pred. No. 1.9e-34;
Matches 107; Conservative 51; Mismatches 98; Indels 23; Gaps 7;
QY 21 APSQVLNFEIDYKEVEEVRGAGFVGVCKAKRAKDAVAKQIESESEKAFIV- 76
DB 83 APAGLQLPQIPHEQLQLEIIGVGFGKYRALWGEVAVKAAALDEKPAVTAEOV 142
QY 77 --ELRQLSRVNHNPVVKLYGACLNLP--VCLVMEYAEAGSLYNNVHGAEPPLPYTTAAHAMS 132
DB 143 CQEARLFGALQHPNIIALRGACLNPHPLCLVMEYARGGALSRLVLAGRRVPPHV- 198
QY 133 WCLQCSQGVAYLHSMQPKALIHRLDKPPNLLV-----AGTVLKI CDFTGACD-IQT 184
DB 199 WAVQVARGMNYLNDAPVPIIHRDLKSNILILEATENNLAADTVLKITDFGLAREWHKT 258
QY 185 HMTNNKGSAAWMAPEVFEFSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMWAV- 243
DB 259 TKMSAAGTYAWMAPEVIRLSFSKSDVMSFGVLLMELLITGVPYREI--DALAVAYGVA 316
QY 244 HNTSRPLINKPLKPIESLMTRCWSKDPQSPMSBIVK 282
```

```
DB 317 MNKLTLPSTPCPEPFARLLLECWDPDPHGRPDRFGSILK 355
RESULT 5
ID M3K9 HUMAN STANDARD; PRT; 394 AA.
AC P80192;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Mitogen-activated protein kinase kinase kinase 9 (EC 2.7.1.1-) (Mixed
DE lineage kinase 1) (Fragment).
GN MAP3K9 OR MLK1 OR PRKE1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon epithelium;
RX MEDLINE=93238756; PubMed=8477742;
RT Dorow D.S., Devereux L., Dietzsch E., de Kretser T.;
RT "Identification of a new family of human epithelial protein kinases
RT containing two leucine/isoleucine-zipper domains.";
RL Eur. J. Biochem. 213:701-710(1993).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN EPITHELIAL TUMOR CELL LINES OF
CC COLONIC, BREAST AND OESOPHAGEAL ORIGIN.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE KINASE SUBFAMILY.
CC PIR; S32467; S32467.
CC PIR; JU0229; JU0229.
CC HSSP; P12931; IFMK.
CC Genew; HGNC:6861; MAP3K9.
CC MTM; 600136; -.
CC InterPro; IPR000719; Euk_pkinase.
CC InterPro; IPR002290; Ser_thr_pkinase.
CC InterPro; IPR001245; Tyr_pkinase.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Euk_pkinase; 1.
CC SMART; SM00219; Tyrc; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase;
KW ATP-binding.
KW NON_TER 1 1
FT DOMAIN 3 271 PROTEIN_KINASE.
FT NP_BIND 9 17 ATP (BY SIMILARITY).
FT BINDING 30 30 ATP (BY SIMILARITY).
FT ACT_SITE 127 127 BY SIMILARITY.
FT DOMAIN 289 310 LEUCINE-ZIPPER 1 (BY SIMILARITY).
FT DOMAIN 324 345 LEUCINE-ZIPPER 2 (BY SIMILARITY).
FT DOMAIN 354 368 ARG/LYS-RICH (BASIC).
FT SEQUENCE 394 AA; 44975 MW; DBE40B7D31047FD8 CRC64;
Query Match 25.9%; Score 418.5; DB 1; Length 394;
Best Local Similarity 36.6%; Pred. No. 2.7e-29;
Matches 98; Conservative 47; Mismatches 100; Indels 23; Gaps 8;
QY 35 BIEVEEVGRGAGFVGVCKAKRAKDAVAK--QIESESEKAFIVELRQ----LSRVNHPN 88
DB 2 ELTLEEIIGVGFGKYVAFWIGDEVAVKAAARHDPDDDISQTIENVVQEAFLKFLKHPN 61
QY 89 IVKLYGACLNLP--VCLVMEYAEAGSLYNNVHGAEPPLPYTTAAHAMSCLQCSQGVAYLHS 146
DB 62 IIALRGVCLKEPNCLVMEYAEAGSLYNNVHGAEPPLPYTTAAHAMSCLQCSQGVAYLHS 177
QY 147 MQPKALIHRLDKPPNLLV-----AGTVLKI CDFTGACD-IQTHMTNNKGSAAWMA 198
DB 118 EAVIPIHRLDKSNILILEATENNLAADTVLKITDFGLAREWHRTYKMSAAGTYAWMA 177
QY 199 EVFEFSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMWAV-HNGTRPFLINKLPK 257
```

```
Db 178 EVIRASMFSGSDVWSGVILMELLTGEVDFRIGD--LRVAVGAMNKLALPISTCPE 235
QY 258 PIESLMTRCWSKDSQSPSPSMEETIKMT 285
Db 236 PFAKLMEDCWNPDHSPRSPFTNILDOLT 263

RESULT 6
KXK2 D1C1D1 STANDARD; PRT; 410 AA.
ID KYK2 D1C1D1
AC P18161;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tyrosine-protein kinase 2 (EC 2.7.1.112) (Fragment).
GN PYKB OR DPYK2.
OS Dictyostelium discoideum (Sjime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90287147; PubMed=1972546;
RA Tan J.L., Spudich J.A.;
RT "Developmentally regulated protein-tyrosine kinase genes in
RT Dictyostelium discoideum."
RL Mol. Cell. Biol. 10:3578-3583(1990).
CC -! CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -! SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M33784; AAA33203.1; -.
CC PIR: B35670; B35670.
CC HSSP: P08631; 1AD5.
CC DictyDb; DD03011; pyKB.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR004040; STY_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00221; STYK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.
FT DOMAIN 1 381 PROTEIN KINASE.
FT NE_BIND 114 122 ATP (BY SIMILARITY).
FT BINDING 135 135 ATP (BY SIMILARITY).
FT ACT_SITE 232 232 BY SIMILARITY.
SQ SEQUENCE 410 AA; 46386 MW; E93918B605B9AC1 CRC64;

Query Match 25.3%; Score 408.5; DB 1; Length 410;
Best Local Similarity 33.8%; Pred. No. 2,1e-28;
Matches 94; Conservative 52; Mismatches 103; Indels 29; Gaps 9;

QY 28 FEE-----IDYKEIEVEEVGGAFGVCAKMKAKVAIKOI-----ESSSERK 72
Db 89 FEEKSLTGEREVIITDNDIOPTOKVGEAPSEVSGMWMGCHVALIKLTIJDEBQFKE 148
QY 73 AFIEHLQRLSRVNHPIVTKLYGACLNPFVCLMEYAEKGSLYNVLH--GAEPUPYYTAH 129
Db 149 RFIIEVGNLKKGNHNTVMFVIGACYKACITTEYMAGSGSLYNIILHNNSSTPRVKYSFPL 208
QY 130 AMSWCLOCSGVAVLHSMQPKALIHRLDKPPNILLVAGGVTLVKICDFGTACD--IOTHTMT 187
```

```
Db 209 VLKATDMALLLHLHSI---TIYHRDLTSNILLDELGNI-KISDGLSAEKREKSMT 264
QY 188 NNKG---SAAMMADEVEEG-SNYSEKCDVPSWGIIIMVEVITRRKPEDEIGAPAFRIMAV 243
Db 265 MTNGGICNPRWRPPELFTKNIIGHYSEKVDVYCFSLVWEIILTGEIIFSDLDG-SQSAQVA 323
QY 244 HNGTRPPLIKNLPKPIESLMTRCWSKDSQSPSPSMEETI 281
Db 324 YAGLRPPIEYCDPELKLTLQCEWADPNDRPPPTIYV 361

RESULT 7
M3KC RAT
ID M3KC RAT STANDARD; PRT; 888 AA.
AC Q63796;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mitogen-activated protein kinase kinase kinase 12 (EC 2.7.1.37)
DE (MAPK-upstream kinase) (MUK).
GN MAP3K12 OR MUK.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9626099; PubMed=8637721;
RA Hirai S., Izawa M., Otsda S., Spyrou G., Ohno S.;
RT "Activation of the JNK pathway by distantly related protein kinases,
RT MEKK and MUK."
RL Oncogene 12:641-650(1996).
CC -! FUNCTION: May be an activator of the JNK/SAPK pathway.
CC phosphorylates beta-casein, histone 1 and myelin basic protein in
CC vitro.
CC -! CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -! COPACTOR: Magnesium.
CC -! SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated (By
CC similarity).
CC -! PM: Autophosphorylated on Ser/Thr. Phosphorylated in cytosol
CC under basal conditions and dephosphorylated when membrane-
CC associated (By similarity).
CC -! SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE KINASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D49785; BAA08621.1; -.
CC HSSP: P12931; IFMK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR004040; STY_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00221; STYK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE; PS00109; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation; Magnesium; Membrane.
FT DOMAIN 158 399 PROTEIN KINASE.
FT NP_BIND 164 172 ATP (BY SIMILARITY).
FT BINDING 185 185 ATP (BY SIMILARITY).
FT ACT_SITE 269 269 BY SIMILARITY.
FT DOMAIN 56 62 POLY-GLY.
FT DOMAIN 668 671 POLY-PRO.
FT DOMAIN 698 701 POLY-PRO.
```


CC STRAIN=ICR X Swiss Webster; TISSUE=Brain;
 RX MEDLINE=96365388; PubMed=8769565;
 RA Blount R., Beaudoin J., Bergeron P., Nadeau A., Grondin G.;
 RT "Cell-specific expression of the zpk gene in adult mouse tissues.";
 RL DNA Cell Biol. 15:631-642(1996).
 RN [3]
 RP PHOSPHORYLATION, AND MUTAGENESIS OF LYS-185 AND GLU-192.
 RX MEDLINE=96279269; PubMed=8663324;
 RA Mata M., Merritt S.E., Fan G., Yu G.G., Holzman L.B.;
 RT "Characterization of dual leucine zipper-bearing kinase, a mixed
 RT kinase present in synaptic terminals whose phosphorylation
 RT state is regulated by membrane depolarization via calcineurin.";
 RL J. Biol. Chem. 271:16888-16896(1996).
 CC - FUNCTION: May be an activator of the JNK/SAPK pathway.
 CC Phosphorylates beta-casein, histone 1 and myelin basic protein in
 CC vitro.
 CC - CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC - COFACTOR: Magnesium.
 CC - SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated.
 CC - TISSUE SPECIFICITY: Expressed in brain, kidney, lung, heart,
 CC testis, gastrointestinal tract, stomach, liver and pancreas.
 CC Within the nervous system, predominantly expressed in neurons and
 CC enriched in synaptic terminals.
 CC - PTM: Autophosphorylated on Ser/Thr. Phosphorylated in cytosol
 CC under basal conditions and dephosphorylated when membrane-
 CC associated.
 CC - SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC MAP KINASE KINASE KINASE SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U14636; AAA57280.1; -;
 DR EMBL: U23789; AAB17123.1; -;
 DR HSSP: P12931; IFMK.
 DR MGD: MGI:1346881; Map3k12.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR004040; STY_pkinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR Pfam: PF00069; pkinase; 1.
 DR ProDom: PD000001; Euk_pkinase; 1.
 DR SMART: SM00221; STYKC1.1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 KM Transferase: Serine/threonine-protein kinase; ATP-binding;
 KM Phosphorylation: Magnesium; Membrane.
 FT DOMAIN 158 399
 FT NP_BIND 164 172
 FT BINDING 185 185
 FT ACT_SITE 269 269
 FT DOMAIN 56 62
 FT DOMAIN 668 671
 FT DOMAIN 698 701
 FT DOMAIN 753 758
 FT MUTAGEN 185 185
 FT MUTAGEN 192 192
 FT CONFLICT 18 18
 FT CONFLICT 28 28
 FT CONFLICT 382 382
 FT CONFLICT 494 495
 FT CONFLICT 517 517
 FT CONFLICT 794 794
 SQ SEQUENCE 888 AA; 96083 MW; CFECCFD34F89ABB CRC64;
 Query Match 24.88; Score 401; DB 1; Length 888;
 Best Local Similarity 33.53; Pred. No. 2,4e-27;
 Matches 88; Conservative 54; Mismatches 99; Indels 22; Gaps 9;

QY 30 EIDYKEIEVEEYVGRGAFVVCARKAKDAVAIKQIESSESRKA FIVELROLSEVNHENI 89
 Db 152 EYFPEBILDLQWVGSGAGQANVFLGRFHGEVAVKVRRLKE-----TDIKRLRLKHNHI 206
 QY 90 VLYGACLNPP--VCLWMEYABEGSLYNVLHGAEPLPYTTAAHMSWCLQCGQVAYLHSM 147
 Db 207 IFKVCVCTQAPCYCTIMEFCAQGLYEVLRAGRPV---TPSLVDWMSWIGAGNNYLLHL 263
 QY 148 GKALIHRLDLKPNLLVLAGGVYVKICDGFACDICTHMTNNK--GSAAMAPVFEBSN 205
 Db 264 K--IIHRLKSPN-WLITTYDDVAKISDGTSKELSDKSTMSRAGVAMWAPVIRNEP 319
 QY 206 YSEKCDVPSWGIILWEVITRRKRPDEIGCPAFRIMWAV-HNGTRPPLIKNPKPIESIMT 264
 Db 320 VSEKVDIMSPGVVLWELLTGEIPYKDVSAA--IIWGSNSLHPVPSSCPDGFKILLR 377
 QY 265 RCWSKDPSCORPSMEIIVKIMTHL 287
 Db 378 QCMNSKPRNRPSPFRQ---ILHL 397
 RESULT 10
 CTR1_ARATH STANDARD; PRT; 821 AA.
 ID CTR1_ARATH
 AC Q05609;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Serine/threonine-protein kinase CTR1 (EC 2.7.1.37).
 GN CTR1 OR AT5G03730 OR F17C15.150.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eusteroideae; Brassicales; Brassicaceae; Arabidopsids.
 ON NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=93161417; PubMed=8431946;
 RA Kieber J.J., Rothenberg M., Roman G., Feldmann K.A., Ecker J.R.;
 RT "CTR1, a negative regulator of the ethylene response pathway in
 RT Arabidopsis, encodes a member of the raf family of protein kinases.";
 RL Cell 72:427-441(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016721; PubMed=11130714;
 RA Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E.,
 RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,
 RA Kohara M., Matsunoto M., Matsuno A., Muraki A., Nakayama S.,
 RA Nakazaki N., Naito K., Okumura S., Shino S., Takeuchi C., Wada T.,
 RA Watanabe A., Yamada W., Yasuda W., Sato S., de la Bastide M.,
 RA Huang B., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,
 RA Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J.,
 RA Stoneking T., Pezin K., Spieth J., Sekhon M., Armstrong J., Becker M.,
 RA Belter E., Cordum H., Cordes M., Courtney L., Courtney W., Dante M.,
 RA Du H., Edwards J., Fryman J., Hakenberg S., Lamat E., Latreille P.,
 RA Leonard S., Meyer R., Mulvaney E., Ozersky P., Riley A., Strommatt C.,
 RA Wagner-McPherson C., Wollam A., Yoakum M., Bell M., Dedhia N.,
 RA Parnell L., Shah R., Rodriguez M., Hoon See L., Vil D., Baker J.,
 RA Kirchhoff K., Toth K., King L., Bahret A., Miller B., Marra M.,
 RA Martensen R., McCombie W.R., Wilson R.K., Murphy G., Bancroft I.,
 RA Volckaert G., Wandut R., Duesterhoeft A., Stekema W., Pohl T.,
 RA Eutlian K.-D., Terryn N., Hartley N., Bent E., Johnson S.,
 RA Langham S.-A., McCullagh B., Robben J., Grymopiez B., Zimmermann W.,
 RA Rampeger U., Wedler H., Balke K., Wedler E., Peters S.,
 RA van Staveren M., Dirkse W., Moolman P., Klein Lankhorst R.,
 RA Weitzneger T., Bothe G., Rose M., Hauf J., Benneiser S., Hempel S.,
 RA Feldpausch M., Lamberth S., Villarroel R., Gelen J., Ardiles W.,
 RA Bents O., Lemcke K., Kolesov G., Mayer K.F.X., Rudd S., Schoof H.,
 RA Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Franz P.F.;
 RT "Sequence and analysis of chromosome 5 of the plant Arabidopsis

```
RT thaliana."
RL Nature 408:823-826 (2000).
CC -!- FUNCTION: ACTS AS A NEGATIVE REGULATOR IN THE ETHYLENE RESPONSE
CC PATHWAY.
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN BOTH SEEDLINGS AND ADULT PLANTS.
CC -!- MISCELLANEOUS: CTR1 MUTANTS DISPLAY ETHYLENE-TREATED PHENOTYPES,
CC RESULTING IN PLANTS WITH SMALL, UNEXPANDED LEAVES AND WHOSE SEED
CC COTYLEDON GROWTH IS IMPAIRED.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MIL/RAF SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L08789; AAA32779.1; -.
CC EMBL; L08790; AAA32780.1; -.
CC EMBL; AL162506; CAB82938.1; -.
CC InterPro; IPR000719; Euk_pkinase.
CC InterPro; IPR004040; STY_pkinase.
CC InterPro; IPR002290; Ser_thr_pkinase.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Euk_pkinase; 1.
CC SMART; SM00221; STYK; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
CC Transferrase; Serine/threonine-protein kinase; ATP-binding.
CC DOMAIN 65 69 POLY-GLY.
CC FT DOMAIN 135 141 POLY-GLY.
CC FT DOMAIN 551 809 PROTEIN_KINASE.
CC FT NP_BIND 557 565 ATP (BY SIMILARITY).
CC FT BINDING 578 578 ATP (BY SIMILARITY).
CC FT ACT_SITE 676 676 BY SIMILARITY.
CC FT MUTAGEN 596 596 E->K: IN CTRL-4; EXHIBITS ETHYLENE-
CC TREATED PHENOTYPE.
CC MUTAGEN 694 694 D->E: IN CTRL-1; EXHIBITS ETHYLENE-
CC TREATED PHENOTYPE.
CC SEQUENCE 821 AA; 90306 MW; 29223D3DCDCC15BC CRC64;
CC -----
Query Match 24.6%; Score 398; DB 1; Length 821;
Best Local Similarity 33.5%; Pred. No. 3.9e-27;
Matches 90; Conservative 60; Mismatches 107; Indels 12; Gaps 8;
QY 30 EIDYKIEVEVVGRGAFGVVCKAKWRAKDVAIK---QIESESER-KAFIVELQQLSRVN 85
DB 545 DPWCDLNTKEKIGAGSFGFVTHRAHGHGSDVAVKILMEQDFHAERVNEFLREVAIMKRLR 604
QY 86 HPNIVKLYGACLPN--VCLVMEYAEGLSYLVNHLGAELPYVYTAAHAMSWCLQCQGVAY 143
DB 605 HPNIVLFMGAVTPPNLSIVTEVLSRGLYRLHLKSGAREQLDERRLSMAVDVAKGMNY 664
QY 144 LHSMPQKALIHRLKLPNLLVAGGTGVLKICDPG-TACDIOHTMNNK--GSAANWAPEV 200
DB 665 LHNRPN-PIVHRLKLPNLLVDKKYTV-KVCDPGLSRKASTFLSKSAAGTPENWAPEV 722
QY 201 FEGSNSEKCDVPSWGIILWEVITRKPKDEIGGPAFRIMVAHVHNGTRPLPKNIKPPIE 260
DB 723 LRDEPNSEKSDVYSFVGIILWEVATLQPGWNL--NPAQVVAAGVFKCKRLRLEIPRNLPQVA 781
QY 261 SLMTRGWSKDPSPORSMEETIVKIMTHLMR 289
DB 782 AITEGCTWEPKRPSEATIMDLRLPLIK 810
RESULT 11
KYK1_DICDI
ID KYK1_DICDI STANDARD; PRT; 1584 AA.
```

```
AC P18160;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Non-receptor tyrosine kinase spore lysis A (EC 2.7.1.112) (Tyrosine-
DE protein kinase 1).
GN PYKA OR SPLA OR DPKYK1.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
CX NCBI_taxid=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JH10;
RX MEDLINE=97053827; PubMed=8898241;
RA Nuckolls G.H., Osherov N., Loomis W.F., Spudich J.A.;
RT "The Dictyostelium dual-specificity kinase spla is essential for
RT spore differentiation."
RL Development 122:3295-3305 (1996).
RN [2]
RP SEQUENCE OF 1248-1584 FROM N.A.
RX MEDLINE=90287147; PubMed=1972546;
RA Tan J.L., Spudich J.A.;
RT "Developmentally regulated protein-tyrosine kinase genes in
RT Dictyostelium discoideum."
RL Mol. Cell. Biol. 10:3578-3583 (1990).
CC -!- FUNCTION: ESSENTIAL FOR SPORE DIFFERENTIATION.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT DEVELOPMENT WITH A PEAK
CC DURING THE MOUND STAGE OF MORPHOGENESIS.
CC -!- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES.
CC -!- SIMILARITY: CONTAINS 1 SAM DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U32174; AAB41125.1; -.
CC EMBL; M33785; AAA33202.1; -.
CC PIR; A35670; A35670.
CC DictyDb; DD03010; pyka.
CC InterPro; IPR000719; Euk_pkinase.
CC InterPro; IPR001660; SAM.
CC InterPro; IPR003878; SPRY domain.
CC InterPro; IPR003877; SPRY receptor.
CC InterPro; IPR004040; STY_pkinase.
CC InterPro; IPR001245; Tyr_pkinase.
CC Pfam; PF00069; pkinase; 1.
CC Pfam; PF00536; SAM; 1.
CC Pfam; PF00622; SPRY; 3.
CC ProDom; PD000001; Euk_pkinase; 1.
CC SMART; SM00454; SAM; 1.
CC SMART; SM00449; SPRY; 3.
CC SMART; SM00221; STYK; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
CC PROSITE; PS00105; SAM_DOMAIN; 1.
CC Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.
FT DOMAIN 908 972 SAM.
FT DOMAIN 403 420 POLY-ASN.
FT DOMAIN 428 435 POLY-THR.
FT DOMAIN 449 480 POLY-ASN.
FT DOMAIN 483 491 POLY-ASN.
FT DOMAIN 494 508 POLY-ASN.
FT DOMAIN 512 532 POLY-ASN.
FT DOMAIN 596 600 POLY-ASN.
FT DOMAIN 808 811 POLY-PHE.
FT DOMAIN 1026 1029 POLY-SER.
```

FT DOMAIN 1195 1210 POLY-ASN.
 FT DOMAIN 1215 1220 POLY-GLN.
 FT DOMAIN 1224 1233 POLY-GLN.
 FT DOMAIN 1266 1274 POLY-PRO.
 FT DOMAIN 1289 1561 PROTEIN KINASE.
 FT NP_BIND 1295 1303 ATP (BY SIMILARITY).
 FT BINDING 1316 1316 ATP (BY SIMILARITY).
 FT ACT SITE 1417 1417 BY SIMILARITY.
 FT CONFLICT 1248 1248 D -> R (IN REF. 2).
 FT CONFLICT 1435 1435 V -> L (IN REF. 2).
 SQ SEQUENCE 1584 AA; 174304 MW; 5D1589458D8E01E3 CMC64;

 Query Match 23.3%; Score 376.5; DB 1; Length 1584;
 Best Local Similarity 32.9%; Pred. No. 6,5e-25;
 Matches 95; Conservative 57; Mismatches 94; Indels 43; Gaps 11;

 QY 30 EIDKEIEVEVGRGAFVGVCAKAKVYAK-----QIESSEKATIVELRQLSRV 84
 DB 1283 EIDNEDEFGOTIGKGFGEYKRGYMETDVAIKIYRDQFKTSSLWMFQNEVGILSKL 1342
 QY 85 NHPNIVKLKGACL---NPVCLVMEYAEAGSL-----YVHLGADPLPYTAAHMSW 133
 DB 1343 RHPVNVQLGACTGAGEDHCHGIVTEWGGSLRQPLTDHFNLL---EQNHFI---RLKL 1395
 QY 134 CLQSGGVAYLHSMOPKXALHRLDKPPNLL-----VAGGYLV--KICDFGTA-- 179
 DB 1396 ALDIAGKNNVYLGWTP--PILHRLDSRRNILLDHNIDPKNPVSSRODIKCKISDFGLSRL 1454
 QY 180 -CDIQTHNTNKGSAAMAPVEFEGSNSEKCDVFSMGIIMEVITRKRPDELGGAFR 238
 DB 1455 KKEQASQMTQGVGCPYAPAEVFGKDSSEKSDVSYGMVLFELLTSDPEQDM-KEMKM 1513
 QY 239 IMWAVHNGTRPPLIKNLKPKIESLMTRCWSQSPQSPMEIEVIMTHL 287
 DB 1514 AHLAYESYRPPILPTTSSKKKEILITQCMDSNPDSRPFKKIIT---VHL 1559

 RESULT 12
 TEC_MOUSE STANDARD; PRT; 630 AA.
 ID TEC_MOUSE
 AC P24604;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tyrosine-protein kinase Tec (EC 2.7.1.112).
 GN TEC.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=93149603; PubMed=7678927;
 RA Mano H., Mano K., Tang B., Koehler M., Yi T., Gilbert D.J.,
 RA Jenkins N.A., Copeland N.G., Ihle J.N.;
 RT "Expression of a novel form of Tec kinase in hematopoietic cells and
 RT mapping of the gene to chromosome 5 near Klt.";
 RT Oncogene 8:417-424(1993).
 RL [2]
 RN SEQUENCE OF 82-630 FROM N.A. (ISOFORM 2).
 RP STRAIN=BALB/C; TISSUE=Liver;
 RX MEDLINE=91133729; PubMed=2284097;
 RA Mano H., Ishikawa F., Nishida J., Hirai H., Takaku F.;
 RT "A novel protein-tyrosine kinase, tec, is preferentially expressed in
 RT liver.";
 RT Oncogene 5:1781-1786(1990).
 RN [3]
 RP SEQUENCE OF 485-553 FROM N.A.
 RX MEDLINE=90152381; PubMed=2482828;
 RA Wilks A.F., Kurban R.R., Hovens C.M., Ralph S.J.;
 RT "The application of the polymerase chain reaction to cloning members
 RT of the protein tyrosine kinase family.";
 RL Gene 85:67-74(1989).

CC - FUNCTION: MAY BE AN IMPORTANT SIGNAL TRANSDUCER FOR CELL DIVISION
 CC AND/OR FOR DIFFERENTIATION IN THE LIVER SYSTEM.
 CC - CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC - SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC - ALTERNATIVE PRODUCTS: 3 isoforms; 1 (shown here) and 2; are
 CC produced by alternative splicing. Isoform 3 may be very rare or
 CC due to a cloning artifact.
 CC - TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN LIVER. EXPRESSION
 CC IS ALSO SEEN IN THE HEMATOPOIETIC CELLS SUCH AS BONE MARROW,
 CC THYMUS AND SPLEEN. LOWER EXPRESSION SEEN IN THE HEART, KIDNEY AND
 CC OVARY.
 CC - SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. BTK
 CC SUBFAMILY.
 CC - SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC - SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC - SIMILARITY: CONTAINS 1 PH DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, S53716; AAA13515.2; -;
 CC EMBL, X55663; CAA39196.1; -;
 CC EMBL, M33427; AAA40018.1; -;
 CC PIR, S13763; S13763.
 CC HSSP, Q06187, 1B55.
 CC GMD, MG1:98662; Tec.
 CC InterPro: IPR001562; BTK.
 CC InterPro: IPR00719; Euk_pkinase.
 CC InterPro: IPR001849; PH.
 CC InterPro: IPR000980; SH2.
 CC InterPro: IPR001452; SH3.
 CC InterPro: IPR001245; Tyr_pkinase.
 CC Pfam, PF00017; SH2; 1.
 CC Pfam, PF00018; SH3; 1.
 CC Pfam, PF00069; pkinase; 1.
 CC Pfam, PF00169; PH; 1.
 CC Pfam, PF00779; BTK; 1.
 CC PRINTS: PR00452; SH3DOMAIN.
 CC PRINTS: PR00402; TECBTKDOMAIN.
 CC PRINTS: PR00109; TYRKINASE.
 CC ProDom, PD000001; Euk_pkinase; 1.
 CC ProDom, PD000066; SH3; 1.
 CC ProDom, PD000093; SH2; 1.
 CC SMART, SM00107; BTK; 1.
 CC SMART, SM00233; PH; 1.
 CC SMART, SM00252; SH2; 1.
 CC SMART, SM00326; SH3; 1.
 CC SMART, SM00219; TYRK; 1.
 CC PROSITE, PS00107; PROTEIN_KINASE_ATP; 1.
 CC PROSITE, PS00109; PROTEIN_KINASE_TYR; 1.
 CC PROSITE, PS50011; PROTEIN_KINASE_DOM; 1.
 CC PROSITE, PS50001; SH2; 1.
 CC PROSITE, PS50002; SH3; 1.
 CC TRANSFERASE, Tyrosine-protein kinase, ATP-binding, SH2 domain;
 KW SH3 domain; Phosphorylation; Alternative splicing.
 FT DOMAIN 4 111
 FT DOMAIN 178 238
 FT DOMAIN 246 344
 FT DOMAIN 369 622
 FT NP_BIND 375 383
 FT BINDING 397 397
 FT ACT SITE 488 488
 FT MOD_RES 518 518
 FT VARSPPLIC 1 94
 VHDANTLYIFAP -> MNVSPVKINFS (IN ISOFORM

```
FT VARSPLIC 224 245 2) MISSING (IN ISOFORM 2).
FT VARSPLIC 604 630 RPEGRSLEDLITIDELVECEFTGR -> ESCLCRVAQD
FT LSKNLIQSRF (IN ISOFORM 3).
FT CONFLICT 535 535 V -> E (IN REF. 3).
FT CONFLICT 550 553 FGLV -> YGIP (IN REF. 3).
FT CONFLICT 611 611 L -> F (IN REF. 2).
SQ SEQUENCE 630 AA; 73426 MW; 262640BE90D4A6D2 CRC64;

Query Match 22.6%; Score 365.5; DB 1; Length 630;
Best Local Similarity 34.3%; Pred. No. 2e-24;
Matches 92; Conservative 49; Mismatches 110; Indels 17; Gaps 9;

QY 30 EIDYKEIEVEEVGGRGAFGVVCKAKWRAK-DVAIKQI-ESESERKAFIVELRQLSRVNH 87
DB 363 EINFSEITFMRELGSGLFGVRLGKWAQYKVAIKAREGAMCEEDFIEAKVMKLT 422
QY 88 NIVKLYGACL--NPVCLVMEYAEGLSYLVHGAEPYPTAAHAMSQCQGGVAYLH 145
DB 423 KLVQLYGVCTQOKPIYIVTFEMERGCLNLFRLQRQ--GHFSRDMLLSMCDVCEGMEYL- 479
QY 146 SMQPKALIHRLDKPPNLLVAGGTVLKICDFGTA----CDIQTHMTNNKGSAAWMAPEVF 201
DB 480 --ERNSTFIHRDLAARNCLVNEAG-VVKVSDFGMARVYLDQYTSSSGAKFPVKWCPPEVF 536
QY 202 EGSNYSEKCDVFSWGIIWEVITR-RKPFDEIGGPAPRIMWAVHNGTRPPLIKNLKPKIE 260
DB 537 NYSRFSKSDVWSFGVLMWEIFTEGRMPFEK--NTNVEVVTWTRGHRHLRPLKATKYLY 594
QY 261 SLMTRCWSKDPQRSPEEIVKINTHLM 288
DB 595 EVMLRCWQERPEGRPSLEDLRTIDELV 622

RESULT 13
TEC_HUMAN STANDARD; PRT; 631 AA.
AC P42680;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tyrosine-protein kinase Tec (EC 2.7.1.112).
GN TEC OR PSCTK4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=95019807; PubMed=7934162;
RA Sato K., Mano H., Ariyama T., Inazawa J., Yazaki Y., Hirai H.;
RT "Molecular cloning and analysis of the human Tec protein-tyrosine
RT kinase.";
RL Leukemia 8:1663-1672(1994).
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- TISSUE SPECIFICITY: HEMATOPOIETIC CELL LINES INCLUDING MYELOID,
CC B-, AND T-CELL LINEAGES.
CC -!- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. BTK
CC SUBFAMILY.
CC -!- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC -!- CAUTION: IT IS UNCERTAIN WHETHER MET-1 IS THE INITIATOR.
CC -!- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
CC WWW="http://www.infobiogen.fr/services/chronocare/Genes/TECID75.html".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
```

```
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D29767; BAA06171.1; -.
DR HSSP; Q06187; 1B55.
DR Genew; HGNC:11719; TEC.
DR MIM; 600583; -.
DR InterPro; IPR001562; BTK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001452; SH3.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF00018; SH3; 1.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00779; BTK; 1.
DR PRINTS; PR00452; SH3DOMAIN.
DR PRINTS; PR00402; TECBKDOMAIN.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Euk_pkinase; 1.
DR ProDom; PD000066; SH3; 1.
DR ProDom; PD000093; SH2; 1.
DR SMART; SM00107; BTK; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00252; SH2; 1.
DR SMART; SM00326; SH3; 1.
DR SMART; SM00219; TyKc; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00001; SH2; 1.
DR PROSITE; PS00002; SH3; 1.
DR PROSITE; PS00003; PH DOMAIN; 1.
KW Transferase; Tyrosine-protein kinase; ATP-binding; SH2 domain;
KW SH3 domain; Phosphorylation.
FT DOMAIN 4 111 PH.
FT DOMAIN 179 239 SH3.
FT DOMAIN 247 345 SH2.
FT DOMAIN 370 623 PROTEIN KINASE.
FT NP_BIND 376 384 ATP (BY SIMILARITY).
FT BINDING 398 398 ATP (BY SIMILARITY).
FT ACT_SITE 489 489 BY SIMILARITY.
FT MOD_RES 519 519 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ SEQUENCE 631 AA; 73629 MW; A55DECAF991A9022 CRC64;
```

```
Query Match 22.6%; Score 365.5; DB 1; Length 631;
Best Local Similarity 34.7%; Pred. No. 2e-24;
Matches 94; Conservative 47; Mismatches 107; Indels 23; Gaps 11;

QY 30 EIDYKEIEVEEVGGRGAFGVVCKAKWRAK-DVAIKQI-ESESERKAFIVELRQLSRVNH 87
DB 364 EINFSEITFMRELGSGLFGVRLGKWAQYKVAIKAREGAMCEEDFIEAKVMKLT 423
QY 88 NIVKLYGACL--NPVCLVMEYAEGLSYLVHGAEPYPTAAHAMSQCQGGVAYLH 145
DB 424 KLVQLYGVCTQOKPIYIVTFEMERGCLNLFRLQRQ--GHFSRDVLLSMCDVCEGMEYL- 480
QY 146 SMQPKALIHRLDKPPNLLVAGGTVLKICDFGTA----CDIQTHMTNNKGSAAWMAPEVF 201
DB 481 --ERNSTFIHRDLAARNCLVSEAG-VVKVSDFGMARVYLDQYTSSSGAKFPVKWCPPEVF 537
QY 202 EGSNYSEKCDVFSWGIIWEVITR-RKPFDEIGGPAPRIMWAVHNGTR----PPLIKNLK 257
DB 538 NYSRFSKSDVWSFGVLMWEVTEGRMPFEKYN--YEVVTWTRGHRHLRPLKATKYLY 593
QY 258 PIESLMTRCWSKDPQRSPEEIVKINTHLM 288
DB 594 -YEVMLRCWQERPEGRPSLEDLRTIDELV 623
```



```

RESULT 14
ITK_HUMAN STANDARD; PRT; 620 AA.
ID ITK_HUMAN
AC 008881;
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tyrosine-protein kinase ITK/TSK (EC 2.7.1.112) (T-cell-specific
kinase) (Tyrosine-protein kinase Ltk) (Kinase EMT).
GN ITK OR LTK OR EMT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93279365; PubMed=8504851;
RA Tanaka N., Asao H., Ohtani K., Nakamura M., Sugamura K.;
RT "A novel human tyrosine kinase gene inducible in T cells by
interleukin 2.";
RL FEBS Lett. 324:1-5(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Thymus; PubMed=8364206;
RA Gibson S., Leung B., Squire J.A., Hill M., Arima N., Goss P.,
Hogg D., Mills G.B.;
RT "Identification, cloning, and characterization of a novel human
T-cell-specific tyrosine kinase located at the hematopoietin complex
on chromosome 5q.";
RL Blood 82:1561-1572(1993).
CC -1- FUNCTION: PLAYS A ROLE IN T CELL PROLIFERATION AND
DIFFERENTIATION.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
tyrosine phosphate.
CC -1- SUBCELLULAR LOCATION: LOCALIZES TO CELL SURFACE RECEPTORS IN THE
PLASMA MEMBRANE AFTER STIMULATION WITH RESPECTIVE RECEPTORS (TCR,
CD28, CD2) IN T-CELLS (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: T-CELL LINES AND NATURAL KILLER CELL LINES.
CC -1- INDUCTION: BY INTERLEUKIN-2.
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. BTK
SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D13720; BAA02873.1; -
DR EMBL; L10717; AAA36748.1; -
DR EMBL; S65186; AAB28072.2; -
DR PIR; S33253; S33253.
DR HSSP; Q03526; 1AMU.
DR Genew; HGNC:6171; ITK.
DR MIM; 186973; -.
DR InterPro; IPR001562; BTK.
DR InterPro; IPR000719; Euk_kinase.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001452; SH3.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF00018; SH3; 1.
DR Pfam; PF00069; PKinase; 1.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00779; BTK; 1.
DR PRINTS; PR00401; SH2DOMAIN.

```

```

DR PRINTS; PR00402; TCBTKDOMAIN.
DR PRINTS; PR00109; TYRKINASE.
DR Prodom; PD000001; Euk_Pkinase; 1.
DR Prodom; PD000066; SH3; 1.
DR Prodom; PD000093; SH2; 1.
DR SMART; SM00107; BTK; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00252; SH2; 1.
DR SMART; SM00326; SH3; 1.
DR SMART; SM00219; TyrcK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS50001; SH2; 1.
DR PROSITE; PS50002; SH3; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
DR Transferase; Tyrosine-protein kinase; Phosphorylation;
KW ATP-binding; SH2 domain; SH3 domain.
KW DOMAIN
FT DOMAIN 171 231 SH3.
FT DOMAIN 239 338 SH2.
FT DOMAIN 363 615 PROTEIN_KINASE.
FT NP_BIND 369 377 ATP (BY SIMILARITY).
FT BINDING 391 391 ATP (BY SIMILARITY).
FT ACT_SITE 482 482 BY SIMILARITY.
FT MOD_RES 512 512 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CONFLICT 331 331 V -> W (IN REF. 2; AAB28072).
SQ SEQUENCE 620 AA; 71831 MW; DAE396BD2309319D CMC64;

Query Match 22.2%; Score 358.5; DB 1; Length 620;
Best Local Similarity 33.5%; Pred. No. 8.2e-24;
Matches 91; Conservative 43; Mismatches 109; Indels 29; Gaps 11;

QY 31 IDKKEIEVEEVGRGAFGVYCKAKWRAD-VAIKOI-ESESRKAFIVELRLQSLVNHVHN 88
DB 358 IDSELTFVQETIGSGQFGVHLGWLNDKVAIKITIRGAMSEEDFIEAEVYMKLSHK 417
QY 89 IYLYIACLN--PVCILMEYAEAGSLVYNLHGAEPDLPYTAHAHNSWCLOCQGVAYLHS 146
DB 418 IYGLYVCELEQAPICLVFPEFMHGCISDLRTQRGL--FAETLLGMCILDVEGNAYL-- 473
QY 147 MQPKALIRHDLKPNLLVAGTGLKICDFG---TACDIQTHMTNNGSAAMAPEVFE 202
DB 474 -EBAVCVHRDLAARN-CLVGENQVIKVSDFGWTFRVLDDQYTSSTGTFKPVVWASPEVFS 531
QY 203 GSNYSKCVFSGNGLIHWVITRRK-PPEITGPAFRIMAVHNGTRPPLIGNLPK--- 258
DB 532 FSTRYSKSDVWSFGVLMWVFESEGIKYENRNS--EVEDISTGFR-----LYKPRLA 583
QY 259 ---IESLMTRCWSKDPSPRSMEEIYKIWTHL 287
DB 584 STHVYQIMNHCKKERPEDRPAPSRLLRLQIAEI 615

RESULT 15
ITK_MOUSE STANDARD; PRT; 625 AA.
ID ITK_MOUSE
AC 003526;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tyrosine-protein kinase ITK/TSK (EC 2.7.1.112) (T-cell-specific
kinase) (Il-2-inducible T-cell kinase) (Kinase EMT) (Kinase TLK).
GN ITK OR TSK OR EMT OR TLK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Thymocytes; PubMed=1280821;
RA Siliciano J.D., Morrow T.A., Desiderio S.V.;

```

RT "itk, a T-cell-specific tyrosine kinase gene inducible by interleukin
RT 2.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:11194-11198(1992).
RN (2)
RC SEQUENCE FROM N.A.
RP TISSUE=Thymocytes;
RX MEDLINE=93133848; PubMed=8421704;
RA Heycock S.D., Berg L.J.;
RT "Developmental regulation of a murine T-cell-specific tyrosine kinase
RL gene, Itk.";
RN Proc. Natl. Acad. Sci. U.S.A. 90:669-673(1993).
RN (3)
RC SEQUENCE FROM N.A.
RP STRAIN=CBA/J; TISSUE=Mast cells;
RX MEDLINE=93236578; PubMed=8476425;
RA Yamada N., Kawakami Y., Kimura H., Fukumachi H., Baier G.,
RA Altman A., Kato T., Inagaki Y., Kawakami T.;
RT "Structure and expression of novel protein-tyrosine kinases, Emb and
RT Ent, in hematopoietic cells.";
RL Biochem. Biophys. Res. Commun. 192:231-240(1993).
RN (4)
RC SEQUENCE FROM N.A.
RP Ogata M., Sawada M., Fujiwara H., Hamaoka T.;
RL Submitted (JAN-1993) to the EMBL/GenBank/DBJ databases.
RN (5)
RC STRUCTURE BY NMR OF 160-236.
RX MEDLINE=97138229; PubMed=8985255;
RA Andreotti A.H., Bunnell S.C., Feng S., Berg L.J., Schreiber S.L.;
RT "Regulatory intramolecular association in a tyrosine kinase of the
RT Tec family.";
RL Nature 385:93-97(1997).
RN (6)
RC CHARACTERIZATION.
RX MEDLINE=95023908; PubMed=7524075;
RA August A., Gibson S., Kawakami Y., Kawakami T., Mills G.B., Dupont B.;
RT "CD28 is associated with and induces the immediate tyrosine
RT phosphorylation and activation of the Tec family kinase ITK/EMT in
RT the human Jurkat leukemic T-cell line.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:9347-9351(1994).
RN (7)
RC CHARACTERIZATION.
RX MEDLINE=97098950; PubMed=8943565;
RA King P.D., Sadra A., Han A., Lu X.-R., Sunder-Plassmann R.,
RA Reinherz E.L., Dupont B.;
RT "CD2 signaling in T cells involves tyrosine phosphorylation and
RT activation of the Tec family kinase, EMT/ITK/TSK.";
RL Int. Immunol. 8:1707-1714(1996).
RN (8)
RC CHARACTERIZATION.
RX MEDLINE=20040393; PubMed=10570288;
RA Ching K.A., Kawakami Y., Kawakami T., Tsoukas C.D.;
RT "Emt/Itk associates with activated TCR complexes: role of the
RT pleckstrin homology domain.";
RL J. Immunol. 163:6006-6013(1999).
CC -!- FUNCTION: MAY PLAY A ROLE IN T-CELL DEVELOPMENT, POTENTIALLY IN
CC THYMIC SELECTION.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBUNIT: LIGATION OF CD2, TCR OR CD28 INDUCES ACTIVATION AND
CC TYROSINE PHOSPHORYLATION OF ITK.
CC -!- SUBCELLULAR LOCATION: LOCALIZES TO CELL SURFACE RECEPTORS IN THE
CC PLASMA MEMBRANE AFTER STIMULATION WITH RESPECTIVE RECEPTORS (TCR,
CC CD28, CD2) IN T-CELLS.
CC -!- TISSUE SPECIFICITY: IS DETECTED IN THE THYMUS, LYMPH NODE AND VERY
CC FAINTLY IN THE SPLEEN, BUT IS NOT DETECTED IN THE LIVER, LUNG,
CC KIDNEY, HEART, BRAIN, INTESTINE OR TESTIS. EXPRESSED IN T-
CC LYMPHOCYTES AND MAST CELLS. IT MAY ALSO BE EXPRESSED IN NATURAL
CC KILLER CELLS.
CC -!- DEVELOPMENTAL STAGE: IS PRESENT IN THE FETAL THYMUS AS EARLY AS
CC DAY 14 OF GESTATION. THE LEVELS ARE 5- TO 10-FOLD HIGHER IN
CC THYMOCYTES THAN IN PERIPHERAL T CELLS, AND INCREASE IN THE THYMUS
CC DURING DEVELOPMENT FROM NEONATE TO ADULT.
CC -!- INDUCTION: BY INTERLEUKIN-2.

CC -!- DOMAIN: THE PH DOMAIN MEDIATES TARGETING OF ITK/TSK AND IS
CC INDISPENSIBLE FOR THE ACTIVATION THROUGH TCR/CD3.
CC -!- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. BTK
CC SUBFAMILY.
CC -!- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L00619; AAA39337.1; -;
CC EMBL; L05631; AAA40518.1; -;
CC EMBL; L10628; -; NOT_ANNOTATED_CDS.
CC EMBL; D14042; BAA03129.1; -;
CC PIR; A43030; A43030.
CC PDB; 1AWJ; 14-JAN-98.
CC MGD; MGI:96621; Itk.
CC InterPro; IPR001562; BTK.
CC InterPro; IPR00779; BTK; 1.
CC InterPro; IPR00779; Euk_pkinase.
CC InterPro; IPR001849; PH.
CC InterPro; IPR00980; SH2.
CC InterPro; IPR001452; SH3.
CC InterPro; IPR001245; Tyr_pkinase.
CC Pfam; PF00017; SH2; 1.
CC Pfam; PF00018; SH3; 1.
CC Pfam; PF00069; pkinase; 1.
CC Pfam; PF00169; PH; 1.
CC Pfam; PF00779; BTK; 1.
CC PRINTS; PR00401; SH2DOMAIN.
CC PRINTS; PR00402; TECTRKDOMAIN.
CC PRINTS; PR00109; TYRKINASE.
CC ProDom; PD000001; Euk_pkinase; 1.
CC ProDom; PD000066; SH3; 1.
CC ProDom; PD000093; SH2; 1.
CC SMART; SM00107; BTK; 1.
CC SMART; SM00233; PH; 1.
CC SMART; SM00252; SH2; 1.
CC SMART; SM00326; SH3; 1.
CC SMART; SM00219; TyrKc; 1.
CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.
CC PROSITE; PS00109; PROTEIN KINASE_TYR; 1.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS50001; SH2; 1.
CC PROSITE; PS50002; SH3; 1.
CC PROSITE; PS50003; PH_DOMAIN; 1.
CC Transferase; Tyrosine-protein kinase; Phosphorylation;
KW ATP-binding; SH2 domain; SH3 domain; 3D-structure.
FT DOMAIN 4 117
FT DOMAIN 177 237
FT DOMAIN 245 343
FT DOMAIN 368 620
FT NP_BIND 374 382
FT BINDING 396 396
FT ACT_SITE 487 487
FT MOD_RES 517 517
FT CONFLICT 82 87
FT CONFLICT 535 535
FT CONFLICT 540 540
SQ SEQUENCE 625 AA; 72291 MW; F7A4A18A8A1AADDC CRC64;
Query Match 22.0%; Score 356; DB 1; Length 625;
Best Local Similarity 32.4%; Pred. No. 1.4e-210;
Matches 93; Conservative 44; Mismatches 110; Indels 40; Gaps 12;
QY 16 GEMIEAPSQLNFEEDIDYKEIEVEEVVGRGAFGVCKAKWRAD-VAIKOI-ESESERKA 73
Db 359 GKWVIOFSELTFFVQEI-----GSGQFGLVHLYWLNKDKVAIKTIQEGAMSEED 407

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 10, 2002, 03:43:10 ; Search time 96 Seconds

(Without alignments)
650.336 Million cell updates/sec

Title: US-09-830-144-2_COPY_1_303

Perfect score: 1615

Sequence: 1 MSTAASASSSSSSSSAGEMIE.....MTIMKRYPPGADEPLQPCQ 303

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1615	100.0	491	4 Q9NZ70	Q9NZ70 homo sapien
2	1609	99.6	606	11 Q9Z3A8	Q9Z3A8 mus musculu
3	1490.5	92.3	616	13 O7J613	O7J613 xenopus lae
4	1428	88.4	478	4 Q9NTR4	Q9NTR4 homo sapien
5	1428	88.4	539	4 Q9NTR1	Q9NTR1 homo sapien
6	1428	88.4	566	4 Q9NTR2	Q9NTR2 homo sapien
7	813	50.3	678	5 Q9VJ06	Q9VJ06 drosophila
8	505	31.3	252	5 Q9VCV0	Q9VCV0 drosophila
9	442	27.4	1066	4 Q9HZN5	Q9HZN5 homo sapien
10	434.5	26.9	411	10 Q9ZQ31	Q9ZQ31 arabidopsis
11	434	26.9	412	10 Q9M085	Q9M085 arabidopsis
12	434	26.9	454	11 Q9ESL3	Q9ESL3 mus musculu
13	434	26.9	802	11 Q9ESL4	Q9ESL4 mus musculu
14	433	26.8	371	13 Q9QZY8	Q9QZY8 brachydanio
15	433	26.8	455	4 Q9HCC4	Q9HCC4 homo sapien
16	433	26.8	800	4 Q9HDD2	Q9HDD2 homo sapien

ALIGNMENTS

17	433	26.8	800	4 Q9HCC5	Q9HCC5 homo sapien
18	433	26.8	800	4 Q9NY12	Q9NY12 homo sapien
19	433	26.8	800	4 Q9NYE9	Q9NYE9 homo sapien
20	432	26.7	564	4 Q9H1Y7	Q9H1Y7 homo sapien
21	430	26.6	422	5 Q23846	Q23846 dictyosteli
22	430	26.6	1001	11 Q8VDD6	Q8VDD6 mus musculu
23	430	26.6	1338	5 Q23927	Q23927 dictyosteli
24	428	26.5	570	4 Q8WMN2	Q8WMN2 homo sapien
25	428	26.5	1036	4 Q8WMN1	Q8WMN1 homo sapien
26	427.5	26.5	391	10 Q22100	Q22100 arabidopsis
27	427.5	26.5	1020	5 Q9W3I3	Q9W3I3 drosophila
28	427.5	26.5	1148	5 Q9SVF6	Q9SVF6 drosophila
29	427.5	26.5	1161	5 Q9SUN8	Q9SUN8 drosophila
30	421.5	26.1	357	10 Q9S1M8	Q9S1M8 arabidopsis
31	421	26.1	850	11 Q9J1J5	Q9J1J5 mus musculu
32	419	25.9	847	4 Q16584	Q16584 homo sapien
33	418.5	25.9	416	10 Q94C42	Q94C42 triticum ae
34	416	25.8	740	5 Q21982	Q21982 caenorhabdi
35	412	25.5	483	10 Q8RY96	Q8RY96 arabidopsis
36	412	25.5	570	10 Q8RWL6	Q8RWL6 arabidopsis
37	411.5	25.5	546	10 Q22558	Q22558 arabidopsis
38	411.5	25.5	553	10 Q81808	Q81808 arabidopsis
39	411	25.4	1030	10 Q9C9U5	Q9C9U5 arabidopsis
40	410	25.4	835	4 Q9Y2V6	Q9Y2V6 homo sapien
41	408	25.3	411	10 Q9AWA6	Q9AWA6 atachis hyp
42	406.5	25.2	370	10 Q9S7D5	Q9S7D5 arabidopsis
43	404	25.0	847	10 Q93XU9	Q93XU9 rosa hybrid
44	401.5	24.9	527	5 Q9B1Z5	Q9B1Z5 dictyosteli
45	401	24.8	859	4 Q8WY25	Q8WY25 homo sapien

RESULT 1

ID	Q9NZ70	PRELIMINARY;	PRT;	491 AA.
AC	Q9NZ70; Q9NTR3;			
DT	01-OCT-2000 (TREMBLrel. 15, Created)			
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	TGF beta-activated kinase splice variant d (D0154G14.1.4)			
DE	(mitogen-activated protein kinase kinase kinase 7 (TGF-beta activated kinase 1d (TAK1)))			
GN	TAK1 OR MAP3K7.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20568765; PubMed=1118615;			
RA	Dempsey C.E., Sakurai H., Sugita T., Guesdon F.;			
RT	"Alternative splicing and gene structure of the transforming growth factor beta-activated kinase 1."			
RT	Biochim. Biophys. Acta 1517:46-52 (2000).			
RL	[2]			
RN	SEQUENCE OF 41-491 FROM N.A.			
RP	Tracey A.;			
RA	Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.			
RL	-1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.			
CC	EMBL; AF218074; AAF27652.1; -.			
DR	EMBL; AL121964; CAB87605.1; -.			
DR	HSSP; P06631; IADS.			
DR	HSSP; P12931; IFMK.			
DR	InterPro; IPR000719; Euk_pkinase.			
DR	InterPro; IPR002290; Ser_thr_pkinase.			
DR	InterPro; IPR004040; Sry_pkinase.			
DR	InterPro; IPR001245; Tyr_pkinase.			
DR	Pfam; PF00069; pkinase_1.			
DR	PRINTS; PR00109; TYRKINASE.			
DR	ProDom; PD000001; Euk_pkinase; 1.			
DR	SMART; SM00221; STYKC; 1.			

```
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 491 AA; 53739 MW; B7D8832E286A99C5 CRC64;

Query Match 100.0%; Score 1615; DB 4; Length 491;
Best Local Similarity 100.0%; Pred. No. 3.5e-148; Indels 0; Gaps 0;
Matches 303; Conservative 0; Mismatches 0;

QY 1 MSTASASSSSSSSSAGEMIEAPSOVLNFEEDYKEIEVEEVVGRGAFGVCKAKWRAKDV 60
DB 1 MSTASASSSSSSSSAGEMIEAPSOVLNFEEDYKEIEVEEVVGRGAFGVCKAKWRAKDV 60

QY 61 AIKQIESESERKAFIVELRQLSRVNHPIVVKLYGACLNPNVCLVMEYAEAGGSLYNVLHGAE 120
DB 61 AIKQIESESERKAFIVELRQLSRVNHPIVVKLYGACLNPNVCLVMEYAEAGGSLYNVLHGAE 120

QY 121 PLPYITAAHAMSWCLOCSOGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFGTAC 180
DB 121 PLPYITAAHAMSWCLOCSOGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFGTAC 180

QY 181 DIQTHMTNKGSAAMWAPVFEFGSNYSEKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIM 240
DB 181 DIQTHMTNKGSAAMWAPVFEFGSNYSEKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIM 240

QY 241 WAVHNGTRPPLIKNLPKPIESLMTRCWSKDPSPQSPMSIEIVKIMTHLMRYFFGADPELOY 300
DB 241 WAVHNGTRPPLIKNLPKPIESLMTRCWSKDPSPQSPMSIEIVKIMTHLMRYFFGADPELOY 300

QY 301 PCQ 303
DB 301 PCQ 303

RESULT 2
Q923A8 PRELIMINARY; PRT; 606 AA.
AC Q923A8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 21, Last annotation update)
DE Unknown (protein for MGC:5989).
GN MAP3K7.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC006665; AA006665.1; -.
DR MGD; MGI:1346877; Map3k7.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.
KW ATP-binding; Transferase.
SQ SEQUENCE 606 AA; 67194 MW; AB8664F389272102 CRC64;

Query Match 99.6%; Score 1609; DB 11; Length 606;
Best Local Similarity 99.7%; Pred. No. 1.8e-147;
Matches 302; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTASASSSSSSSSAGEMIEAPSOVLNFEEDYKEIEVEEVVGRGAFGVCKAKWRAKDV 60
DB 1 MSTASASSSSSSSSAGEMIEAPSOVLNFEEDYKEIEVEEVVGRGAFGVCKAKWRAKDV 60

QY 61 AIKQIESESERKAFIVELRQLSRVNHPIVVKLYGACLNPNVCLVMEYAEAGGSLYNVLHGAE 120
```

```
DB 61 AIKQIESESERKAFIVELRQLSRVNHPIVVKLYGACLNPNVCLVMEYAEAGGSLYNVLHGAE 120
QY 121 PLPYITAAHAMSWCLOCSOGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFGTAC 180
DB 121 PLPYITAAHAMSWCLOCSOGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFGTAC 180

QY 181 DIQTHMTNKGSAAMWAPVFEFGSNYSEKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIM 240
DB 181 DIQTHMTNKGSAAMWAPVFEFGSNYSEKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIM 240

QY 241 WAVHNGTRPPLIKNLPKPIESLMTRCWSKDPSPQSPMSIEIVKIMTHLMRYFFGADPELOY 300
DB 241 WAVHNGTRPPLIKNLPKPIESLMTRCWSKDPSPQSPMSIEIVKIMTHLMRYFFGADPELOY 300

QY 301 PCQ 303
DB 301 PCQ 303

RESULT 3
Q73613 PRELIMINARY; PRT; 616 AA.
AC Q73613;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE TAKI.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=98130593; PubMed=9463380;
RA Shibuya H., Iwata H., Masuyama N., Gotoh Y., Yamaguchi K., Irie K.,
RA Matsumoto K., Nishida E., Ueno N.;
RT "Role of TAKI and TAB1 in BMP signaling in early Xenopus
RT development.";
RL EMBL J. 17:1019-1028 (1998).
CC -/- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; U92030; AAC14008.1; -.
DR HSP; P12931; 1PMK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00221; STYKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 616 AA; 68464 MW; 493AD2A05ADC38B6 CRC64;

Query Match 92.3%; Score 1490.5; DB 13; Length 616;
Best Local Similarity 94.9%; Pred. No. 5.7e-136;
Matches 277; Conservative 6; Mismatches 8; Indels 1; Gaps 1;

QY 12 SSSAGEMIEAPSOVLNFEEDYKEIEVEEVVGRGAFGVCKAKWRAKDVAIKQIESESER 71
DB 2 SATSAEMIETP-PVLNFEEDYKEIEVEEVVGRGTFGVCKAKWRAKDVAIKQIESESER 60

QY 72 KAFIVELRQLSRVNHPIVVKLYGACLNPNVCLVMEYAEAGGSLYNVLHGAEPLPYITAAHAM 131
DB 61 KAFIVELRQLSRVNHPIVVKLYGACLNPNVCLVMEYAEAGGSLYNVLHGAEPLPYITAAHAM 120

QY 132 SWCLOCSOGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFGTACDIQTHMTNKG 191
DB 121 SWCLOCSOGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFGTACDIQTHMTNKG 180

QY 192 SAAMWAPVFEFGSNYSEKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMAVHNGTRPPL 251
```

```
Db 181 SAAMMAPEVEGSGNSYSEKCDVFSWGIIIMVITRRKPFDEIGGAPAFRIMAVHNGTRPPL 240
QY 252 IKULPRIESLMTRCWSKDPSPSMEEIYKIMTHLMRYPPGADPLQYPCQ 303
Db 241 IKULPRIESLMTRCWSKDPSPSMEEIYKIMTHLMQYPPGADVSLQYPCQ 292

RESULT 4
Q9NTR4 PRELIMINARY; PRT; 478 AA.
ID Q9NTR4 AC Q9NTR4;
AC Q9NTR4;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE D154G14.1.3 (Mitogen-activated protein kinase kinase 7
DE (TGF-beta activated kinase 1c (TAK1))) (Fragment).
GN MAP3K7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tracey A.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AL121964; CAB87604.1; -.
DR HSSP; P12931; 1PMK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00221; STYKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Kinase; Serine/threonine-protein kinase.
FT NON_TER 1
SQ SEQUENCE 478 AA; 52482 MW; 177CC8CFA8D8BF8 CRC64;

Query Match 88.4%; Score 1428; DB 4; Length 478;
Best Local Similarity 100.0%; Pred. No. 4, 8e-130; Indels 0; Gaps 0;
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 VVGRGAFGVCCAKWRADVAIKQISESESKAFIVELRQLSRVNHPIVLYGACINPV 100
Db 1 VVGRGAFGVCCAKWRADVAIKQISESESKAFIVELRQLSRVNHPIVLYGACINPV 60
QY 101 CLVMEYABGSGLYNVLHGAEPYPYTTAAHAMSWCLOCSQGVAYIHSNOPKALIHRLDKP 160
Db 61 CLVMEYABGSGLYNVLHGAEPYPYTTAAHAMSWCLOCSQGVAYIHSNOPKALIHRLDKP 120
QY 161 NLLVAGGTIVLKICDPFGACDIQTHMTNNGSAAMMAPEVEGSGNSYSEKCDVFSWGIIIM 220
Db 121 NLLVAGGTIVLKICDPFGACDIQTHMTNNGSAAMMAPEVEGSGNSYSEKCDVFSWGIIIM 180
QY 221 EVITRRKPFDEIGGAPAFRIMAVHNGTRPPLIKULPKPIESLMTRCWSKDPSPSMEEI 280
Db 181 EVITRRKPFDEIGGAPAFRIMAVHNGTRPPLIKULPKPIESLMTRCWSKDPSPSMEEI 240
QY 281 VKIMTHLMRYPPGADPLQYPCQ 303
Db 241 VKIMTHLMRYPPGADPLQYPCQ 263

RESULT 5
Q9NTR1 PRELIMINARY; PRT; 539 AA.
AC Q9NTR1;
RN [1]
RP SEQUENCE FROM N.A.
```

```
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE D154G14.1.1 (Mitogen-activated protein kinase kinase 7
DE (TGF-beta activated kinase 1a (TAK1))) (Fragment).
GN MAP3K7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

Query Match 88.4%; Score 1428; DB 4; Length 539;
Best Local Similarity 100.0%; Pred. No. 5, 6e-130; Indels 0; Gaps 0;
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 VVGRGAFGVCCAKWRADVAIKQISESESKAFIVELRQLSRVNHPIVLYGACINPV 100
Db 1 VVGRGAFGVCCAKWRADVAIKQISESESKAFIVELRQLSRVNHPIVLYGACINPV 60
QY 101 CLVMEYABGSGLYNVLHGAEPYPYTTAAHAMSWCLOCSQGVAYIHSNOPKALIHRLDKP 160
Db 61 CLVMEYABGSGLYNVLHGAEPYPYTTAAHAMSWCLOCSQGVAYIHSNOPKALIHRLDKP 120
QY 161 NLLVAGGTIVLKICDPFGACDIQTHMTNNGSAAMMAPEVEGSGNSYSEKCDVFSWGIIIM 220
Db 121 NLLVAGGTIVLKICDPFGACDIQTHMTNNGSAAMMAPEVEGSGNSYSEKCDVFSWGIIIM 180
QY 221 EVITRRKPFDEIGGAPAFRIMAVHNGTRPPLIKULPKPIESLMTRCWSKDPSPSMEEI 280
Db 181 EVITRRKPFDEIGGAPAFRIMAVHNGTRPPLIKULPKPIESLMTRCWSKDPSPSMEEI 240
QY 281 VKIMTHLMRYPPGADPLQYPCQ 303
Db 241 VKIMTHLMRYPPGADPLQYPCQ 263

RESULT 6
Q9NTR2 PRELIMINARY; PRT; 566 AA.
ID Q9NTR2 AC Q9NTR2;
AC Q9NTR2;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE D154G14.1.2 (Mitogen-activated protein kinase kinase 7
DE (TGF-beta activated kinase 1b (TAK1))) (Fragment).
GN MAP3K7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
```

Tracey A.;
 RA Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AL121964; CAB87606.1; -;
 DR HSSP; P12931; 1FMK.
 DR InterPro; IPR000719; Euk_pkinase.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR InterPro; IPR004040; STY_pkinase.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Euk_pkinase; 1.
 DR SMART; SM00221; STYKC; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW Kinase; Serine/threonine-protein kinase.
 FT NON_TER 1
 SQ SEQUENCE 566 AA; 62972 MW; 32CDAC1211B200CF CRC64;
 Query Match 88.4%; Score 1428; DB 4; Length 566;
 Best Local Similarity 100.0%; Pred. No. 6e-130;
 Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 41 VVGRGAFGVVCKAKWRAKDAIVAIAKQIESERKAFIVELRQLSRVNHNPVIVKLYGACLNVPV 100
 DB 1 VVGRGAFGVVCKAKWRAKDAIVAIAKQIESERKAFIVELRQLSRVNHNPVIVKLYGACLNVPV 60
 QY 101 CLVNEYAEGSLNVNLHGAEPLPYVYTAAHAMSWCLOCSQGVAYLHSMQKALIHRLDKPP 160
 DB 61 CLVNEYAEGSLNVNLHGAEPLPYVYTAAHAMSWCLOCSQGVAYLHSMQKALIHRLDKPP 120
 QY 161 NLLLVAGGTVLKICDFGTACDIOTHTNNKGSAAWVAPEVFGNSYSEKCDVFSWGIIILW 220
 DB 121 NLLLVAGGTVLKICDFGTACDIOTHTNNKGSAAWVAPEVFGNSYSEKCDVFSWGIIILW 180
 QY 221 EVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNLKPESLMTRCWSKDPORPSMEI 280
 DB 181 EVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNLKPESLMTRCWSKDPORPSMEI 240
 QY 281 VKIMTHLMRYFPGADEPLQVPCQ 303
 DB 241 VKIMTHLMRYFPGADEPLQVPCQ 263
 RESULT 7
 Q9V3Q6 PRELIMINARY; PRT; 678 AA.
 AC Q9V3Q6;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Putative TAK1 protein (CGI388 protein) (LD42274P).
 GN TAK1 OR CGI388 OR CGI8492.
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., M.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 [2]
 SEQUENCE FROM N.A.
 RP Takatsu Y., Nakamura M., Stapleton M., Danos M., Matsumoto M.,
 RA O'Connor M.B., Shibuya H., Ueno N.;
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nuncio J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Nuncio J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AE003571; AAF50895.1; -;
 DR EMBL; AF199466; AAF06815.1; -;
 DR EMBL; AY051953; AAK93377.1; -;
 DR HSSP; P08631; IAD5.
 DR FlyBase; FBgn0026323; Tak1.
 DR InterPro; IPR000719; Euk_pkinase.
 DR InterPro; IPR001990; Granin.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR InterPro; IPR004040; STY_pkinase.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Euk_pkinase; 1.
 DR SMART; SM00221; STYKC; 1.
 DR PROSITE; PS00422; GRANINS_1; UNKNOWN_1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Hypothetical protein; Kinase;
 KW Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 678 AA; 75674 MW; 87EBA80CDB8CDE45 CRC64;

Query Match 50.3%; Score 813; DB 5; Length 678;
 Best Local Similarity 54.0%; Pred. No. 3.8e-70;
 Matches 154; Conservative 45; Mismatches 82; Indels 4; Gaps 3;
 QY 18 MIEAPSQLVNPFEIDYKEIEVEEVVGAGVGVCKAKWRAKDAIVAIAKQIESERKAFIVE 77
 DB 1 MATASLDALQAAYVDFSEITLREKVGHSYGVVCKAVWRDKLVAVKEFFASAQKQIEKE 60


```
QY 78 LRQSRVNHPIVLYG--ACINPVCUMVEYABEGSLVYNLHGAEPILPYTAHAMSACL 135
DB 61 VKQISRVKHPHPIALHGISSYQATYILIMEFAEGSLHNFHAG-KVPAFVSLHAMSAR 119
QY 136 QCSGVVYVLSMOPKALIHRLDKPNNLLVAGGTVLKICDFTACDIQTHMTNKSAAW 195
DB 120 QCAAGLIVLHAMPKPLIHNDVKKPLNLLLNKGRNLKICDFTVADKSTMTNNKSGAAW 179
QY 196 MAPEVFGSNVSEKCDVFSWGIIIMEVITRRKPEDEIGGAFRIMAVHNGTRPPLIKNL 255
DB 180 MAPEVFGSNVSEKCDIFSVNVLVMEVLSRKQPKGIDN-AYTLQMKIYKGEHPPLLTTC 238
QY 256 PKPIESLMTKCSQDPSQPSMEVITKIMTHLWRYFPGADEPLQY 300
DB 239 PKRIEDLMTACWKVTPBDRPSMQYIVGMHEIVKDYTGADKALEY 283

RESULT 8
Q9VCV0 PRELIMINARY; PRT; 252 AA.
ID Q9VCV0
AC Q9VCV0;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE CG4803 protein.
GN CG4803.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
ON NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gccayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champagne M., Pfeiffer B.D.,
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Adair J.F., Agbayan A., An H.-U., Andrews-Pfannkuch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howard T.J., Wei M.-H., Ibegum C.,
RA Jajal M., Kalush F., Karpen G.H., Ke Z., Kemison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mekulov G., Milshina N.V., Mobarry C., Morris J., Moshireit A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RL "The genome sequence of Drosophila melanogaster";
Science 287:2185-2195(2000).
```

```
CC -I- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AE003740; AAF56055.1; -.
DR FlyBase; FBgn0039015; CG4803.
DR InterPro; IPR000719; Euk_Pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_Pkinase; 1.
DR PROSITE; PS00107; PROTEIN KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE_ST; 1.
KW ATP-binding; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 252 AA; 28853 MW; 8E006F2EAB8BD17 CRC64;

Query Match 31.3%; Score 505; DB 5; Length 252;
Best Local Similarity 40.6%; Pred. No. 8.2e-41;
Matches 99; Conservative 48; Mismatches 71; Indels 26; Gaps 4;

QY 29 EIDDKYKEIFEEVVRGAFGVYCAKMAKVAIKQISESEKAFIVELPQLSRVHNP 88
DB 6 EGVPEEIIQTKELIGTGYGSGYRAVMNRRIALKRIREGCEDKKIEREIVQLTKASHVN 65
QY 89 IYKLGACILNPPVC--LVMEYABEGSLVYNLHGAEPILPYTAHAMSACLQCSGVAVYTHS 146
DB 66 IVELYGTSRHEGCALLMEFVDGSLSSFLH-AKSKPEYSHAHAFNWAHQIAQGIAYLHG 124
QY 147 MQPKALIHRLDKPNNLLVAGGTVLKICDFTACDIQTHMTNKSAAWAPPE----- 199
DB 125 MQPKAVIHRDILKPLNLLLNKGRNLKICDFTVADKSTMTNNKSGAAWPEVRELFDF 184
QY 200 -----VFSGSNVSEKCDVFSWGIIIMEVITRRKPEDEIGGAFRIMAVH 244
DB 185 KSNRIIINOPTGFOVKVLOGNKPDEKCDVYSWATIFWEILSRKEPEQY-NTLFELYMAIN 243
QY 245 NGTR 248
DB 244 EGKR 247

RESULT 9
Q9H2N5 PRELIMINARY; PRT; 1066 AA.
ID Q9H2N5
AC Q9H2N5;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Mixed lineage kinase MLK1 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
ON NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA McNeae J.U., Dower S.K., Guesdon F.;
RL "cDNA sequence and gene organisation of mixed lineage kinase 1.";
RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
CC -I- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
DR EMBL; AF251442; AAG44591.1; -.
DR HSRF; P29355; ISEM.
DR InterPro; IPR000719; Euk_Pkinase.
DR InterPro; IPR001452; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PR00452; SH3DOMAIN.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Euk_Pkinase; 1.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00326; SH3; 1.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS00107; PROTEIN KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN KINASE_DOM; 1.
```


SEQUENCE 412 AA; 46083 MW; 644F35A90210D488 CRC64;
Query Match 26.9%; Score 434; DB 10; Length 412;
Best Local Similarity 36.8%; Pred. No. 1.2e-33;
Matches 103; Conservative 56; Mismatches 89; Indels 32; Gaps 11;
QY 25 VLNEE--IDYKEIEVEEVGRGAFGVCCAKMRAKVAIKQIE-----SESERKAF 74
DB 118 LVNVEEHTIDIRKLHMGPAFAQGAFGKLYRGTYNGEDVAIKLERSDSNDEKAQALEQCF 177
QY 75 IVEIKOLSRVNHPIVLYGACLNVP--CLVMEYAEGGSLYNVL----HGAELPYTYTAA 128
DB 178 QQEVSMIAFLGHPIVPIFGCIKPMWCVITEYAKGSVRQFTFKRQNRAPVPLKT---- 233
QY 129 HANWSCLOCSQGVAYVLSMOPKALIHRLDKPNNLLVAGGTVLKICDFGTA-CDIOTH-M 186
DB 234 -AVWQALDVARGMAYVHE---RNFIRHDLKSDNLLISADSI-KIADFGYARIEVQTEGM 288
QY 187 TNNKGAAMAAPEVEEGSNYSEKCVFSWGIIMVEVITRRKPEDEIGG--PAFRIMWAVH 244
DB 289 TPETGYRWMAPEMIQHRPPTQKVDVYSPGIVLWELITGLLPQNMTAVOAAFAV---VN 345
QY 245 NGTPPPLIKNLPRKPIESIMTRCWSKDSQRPSEIYKIM 284
DB 346 RGVPRPTVADCLPVLGEIMTRCWDADPEVRPCFAEIVNLL 385
RESULT 12
Q9ESL3 PRELIMINARY; PRT; 454 AA.
AC Q9ESL3; 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE MLTK-beta.
GN ZAK OR MLTK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21264927; Pubmed=11042189;
RA Gotoh I.; Adachi M.; Nishida E.;
RT "Identification and Characterization of a Novel MAP Kinase Kinase
RT Kinase, MLTK.";
RL J. Biol. Chem. 276:4276-4286(2001).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB049732; BAB16443.1; -.
DR HSSP; P12931; IFMK.
DR MGD; MGI:1931274; ZAK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR004040; STY_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR Prodom; PD000001; Euk_pkinase; 1.
DR SMART; SM00221; STYKc; 1.
DR SMART; SM00219; S_TKc; 1.
DR SMART; SM00219; TYKc; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 454 AA; 51366 MW; 35C2FC0D729D9395 CRC64;
Query Match 26.9%; Score 434; DB 11; Length 454;
Best Local Similarity 36.7%; Pred. No. 1.4e-33;
Matches 99; Conservative 53; Mismatches 90; Indels 28; Gaps 11;
QY 27 NFEIDYKEIEVEEVGRGAFGVCCAKM--RAKDVAIK--QIESSEKRAFIIVEIROL 81
DB 7 SFVQIKFDLDLQFFENCGGSGFSYVRAKMTISQDKEVAVKLLKIEKAE-----IL 57

SRVNHPIVLYGACLNIP--VCLVMEYAEGGSLYNVLHG--AEPLPYTYTAHANSWCLOQC 137
DB 58 SVLSHRNIIIQFGVILPEPNVIGVIEYASLSGLYDINSNSEEN---DMEHIMTWADTV 114
QY 138 SQGVAYVLSMOPKALIHRLDKPNNLLVAGGTVLKICDFGTA-CDIOTHMNNK--GSAAW 195
DB 115 AKGMHYLHMEAFVVKVILHDLKSRNVVIAADG-VLKICDFG-ASRPHNHTHMSLVGTFFP 172
QY 196 MAPEVEEGSNYSEKCVFSWGIIMVEVITRRKPEDEIGGPAFRIMW-AVNHGTPPPLIKN 254
DB 173 MAPEVIGSLPVSSETDITYSGVILWEMLTREVFPKGLBG--LQVAMLVVEKNERLLTIPS 230
QY 255 LKPIESIMTRCWSKDSQRPSEIYKIM 284
DB 231 CPRSFAPLHQCWEADAKKRPBFKOIISIL 260
RESULT 13
Q9ESL4 PRELIMINARY; PRT; 802 AA.
AC Q9ESL4; 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE MLTK alpha.
GN ZAK OR MLTK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21264927; Pubmed=11042189;
RA Gotoh I.; Adachi M.; Nishida E.;
RT "Identification and Characterization of a Novel MAP Kinase Kinase
RT Kinase, MLTK.";
RL J. Biol. Chem. 276:4276-4286(2001).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB049731; BAB16442.1; -.
DR HSSP; P12931; IFMK.
DR MGD; MGI:1931274; ZAK.
DR InterPro; IPR000194; Arpase a/bcentre.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR001660; SAM.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR004040; STY_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00516; SAM; 1.
DR PRINTS; PR00109; TYRKINASE.
DR Prodom; PD000001; Euk_pkinase; 1.
DR SMART; SM00454; SAM; 1.
DR SMART; SM00221; STYKc; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TYKc; 1.
DR PROSITE; PS00152; ARPASE_ALPHA_BETA; UNKNOWN_1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 802 AA; 91719 MW; DA3IDF8F312A43CC CRC64;
Query Match 26.9%; Score 434; DB 11; Length 802;
Best Local Similarity 36.7%; Pred. No. 3e-33;
Matches 99; Conservative 53; Mismatches 90; Indels 28; Gaps 11;
QY 27 NFEIDYKEIEVEEVGRGAFGVCCAKM--RAKDVAIK--QIESSEKRAFIIVEIROL 81
DB 7 SFVQIKFDLDLQFFENCGGSGFSYVRAKMTISQDKEVAVKLLKIEKAE-----IL 57
QY 82 SFVNHPIVLYGACLNIP--VCLVMEYAEGGSLYNVLHG--AEPLPYTYTAHANSWCLOQC 137
DB 58 SVLSHRNIIIQFGVILPEPNVIGVIEYASLSGLYDINSNSEEM---DMEHIMTWADTV 114

```
QY 138 SOGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFGTACDIQTHMTNNK--GSAAW 195
Db 115 AKGMHYLHWEAPVKVIHRLDKSRNVVIAADG-VLKICDFG-ASRFHNHTTHMSLVGTFPW 172
QY 196 MAPEVPEGNSYSEKCDVFSWGIIWEVITRRKPFDEIGGPAFRIMW-AVHNGTRPPLIKN 254
Db 173 MAPEVIQSLPVSETCDTYSYGVVLWEMLTREVPFKGLEG--LQVAWLVEKNERLITPSS 230
QY 255 LPKPIESLMTRCWSKDPSPRSMEEIVKIM 284
Db 231 CPRSFAELLHQWEADAKRPSFKIISIL 260

RESULT 14
Q90ZY8 PRELIMINARY; PRT; 371 AA.
AC Q90ZY8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Protein kinase Npk.
OS Brachydanio rerio (Zebrafish) (Zebra daniel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN SEQUENCE FROM N.A.
RA Chou C.-M., Lee I.-L., Leu J.-H., Huang C.-J.;
RT "A novel protein kinase, zNPK, from the zebrafish.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF265343; AAKS2416.1; -.
DR InterPro; IPR000719; Ser_thr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.
KW ATP-binding; Kinase; Transferase.
SQ SEQUENCE 371 AA; 42456 MW; 9B918B8A8B20D296 CRC64;

Query Match 26.8%; Score 433; DB 13; Length 371;
Best Local Similarity 37.1%; Pred. No. 1.3e-33;
Matches 101; Conservative 50; Mismatches 97; Indels 24; Gaps 10;

QY 20 EAPSOVLNFEEDYKIEVEEVVGRGAFGVCKAKW--RAKDVAIK---QIESESERKAF 74
Db 31 ENSSLASFPQIPFDDIRFVENCGGSGFSGVYRAHWVPQDKEVAVKLLKIDAEAE--- 86
QY 75 IVELRQLSRVNHNPVVKLYGACLNVP--CLVMEYAEGLSLYNVLHGAPLPYYTAAHAMS 132
Db 87 -----ILSVLSHKNITQFYGAILEAPNDGIVTEYASRGSLEYLSADS--EEMDMQVMT 140
QY 133 WCLQCSQGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFGTACDIQTHMTNNK-- 190
Db 141 WAMEIAGHYLHAEAPLVIIHRLDKSRNVVITA-DNVLKICDFG-ASKWSHHTHMSLV 198
QY 191 GSAAMMAPEVFEFGNSYSEKCDVFSWGIIWEVITRRKPFDEIGGPAFRIMW-AVHNGTRP 249
Db 199 GTFPWWAPEVIQSLPVSETCDTYSYGVVLWEMLTREVPFK--GFEGQLQVAWLVEKHERP 256
QY 250 PLIKNPKPIESLMTRCWSKDPSPRSMEEIV 281
Db 257 TIPSSCPASFADLMRRRCWNAEPKRPQFKQIL 288

RESULT 15
Q9HCC4 PRELIMINARY; PRT; 455 AA.
AC Q9HCC4;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
```

```
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE MLTK-beta (Similar to sterile-alpha motif and leucine zipper
DE containing kinase AZK) (Mixed lineage kinase) (Mixed lineage
DE kinase-related kinase MRK-beta).
GN MLTK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1].
RP SEQUENCE FROM N.A.
RA Gotoh I., Adachi M., Nishida E.;
RT "Identification and Characterization of a Novel MAP Kinase Kinase
RT Kinase, MLTK.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN [2].
RP SEQUENCE FROM N.A.
RC TISSUE=COLON;
RA Strausberg R.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [3].
RP SEQUENCE FROM N.A.
RA Acton S.;
RT "MLK-mixed lineage kinase.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [4].
RP SEQUENCE FROM N.A.
RX MEDLINE=21950776; PubMed=11836244;
RA Gross E.A., Callow M.G., Waldbaum L., Thomas S., Ruggieri R.;
RT "MRK, a Mixed Lineage Kinase-related Molecule That Plays a Role in
RT gamma-Radiation-induced Cell Cycle Arrest.";
RL J. Biol. Chem. 277:13873-13882(2002).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB049734; BAB16445.1; -.
DR EMBL; BC001401; AAOH01401.1; -.
DR EMBL; AF325454; AAK11615.1; -.
DR EMBL; AF480462; AAL85892.1; -.
DR HSP; P12931; 1FMK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR022290; Ser_thr_kinase.
DR InterPro; IPR004040; Tyr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00221; STYK; 1.
DR SMART; SM00219; TYRK; 1.
DR SMART; SM00220; S_TK; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 455 AA; 51582 MW; E87DB84A4D58B752 CRC64;

Query Match 26.8%; Score 433; DB 4; Length 455;
Best Local Similarity 36.7%; Pred. No. 1.7e-33;
Matches 99; Conservative 53; Mismatches 90; Indels 28; Gaps 11;

QY 27 NFEIDYKTEVEEVVGRGAFGVCKAKW--RAKOVAIK---QIESESERKAFIVELROL 81
Db 7 SFQIKFDDQLQFPENC GGSGFSGVYRAKMTSQDKEVAVKLLKIEKAE-----IL 57
QY 82 SRVNHNPVVKLYGACLNVP--VCLVMEYAEGLSLYNVLHG--AEPLPYTAAHAMSCLQC 137
Db 58 SVLSHRNIIQFYGVILEPPNYGIVTEYASLGSLYDYNINRSEEM---DMDHIMTWATDV 114
QY 138 SOGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFGTACDIQTHMTNNK--GSAAW 195
Db 115 AKGMHYLHWEAPVKVIHRLDKSRNVVIAADG-VLKICDFG-ASRFHNHTTHMSLVGTFPW 172
QY 196 MAPEVFEFGNSYSEKCDVFSWGIIWEVITRRKPFDEIGGPAFRIMW-AVHNGTRPPLIKN 254
Db 173 MAPEVIQSLPVSETCDTYSYGVVLWEMLTREVPFKGLEG--LQVAWLVEKNERLITPSS 230
```

```
Qy      255 LPKPIESIMTRCNSKIDPSQRPSEIIVKIM 284
      | : : : | : : : | :
Db      231 CPRSFAELLHQCNWADAKRPSFKQIISIL 260
```

Search completed: December 10, 2002, 03:50:38
Job time : 98 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 10, 2002, 03:46:30 ; Search time 38 Seconds
(without alignments)
234.609 Million cell updates/sec

Title: US-09-830-144-2_COPY_1_303
Perfect score: 1615
Sequence: 1 MSTASASSSSSSSAGEMIE.....MTHLMRYFPQADEPLQYPCQ 303

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1615	100.0	579	US-09-529-279-4	Sequence 4, Appli
2	1615	100.0	590	US-09-529-279-15	Sequence 15, Appli
3	433	26.8	455	US-09-221-235-5	Sequence 5, Appli
4	433	26.8	455	US-09-221-928-5	Sequence 5, Appli
5	433	26.8	455	US-09-221-527-5	Sequence 5, Appli
6	433	26.8	455	US-09-221-236-5	Sequence 5, Appli
7	433	26.8	455	US-09-221-416-5	Sequence 5, Appli
8	433	26.8	455	US-09-221-245-5	Sequence 5, Appli
9	433	26.8	455	US-09-163-115-5	Sequence 5, Appli
10	433	26.8	455	US-09-221-528-5	Sequence 5, Appli
11	433	26.8	455	US-09-593-553-5	Sequence 5, Appli
12	433	26.8	455	US-09-221-237-5	Sequence 5, Appli
13	410	25.4	835	US-09-291-839-2	Sequence 2, Appli
14	401	24.8	668	US-08-205-018-2	Sequence 2, Appli
15	401	24.8	859	US-08-395-580-2	Sequence 2, Appli
16	401	24.8	859	PCT-US95-02792-2	Sequence 2, Appli
17	398	24.6	821	US-07-928-464-2	Sequence 2, Appli
18	398	24.6	821	US-08-003-311B-2	Sequence 2, Appli
19	398	24.6	821	US-08-261-432-2	Sequence 2, Appli
20	398	24.6	821	PCT-US93-07347-2	Sequence 2, Appli
21	393	24.3	263	US-09-035-706-5	Sequence 5, Appli
22	393	24.3	263	US-08-955-841-5	Sequence 5, Appli
23	393	24.3	263	US-09-390-425-5	Sequence 5, Appli
24	393	24.3	263	US-09-566-906-5	Sequence 5, Appli
25	376.5	23.3	1584	US-09-457-040B-27	Sequence 27, Appli
26	363.5	22.5	527	US-08-426-509A-10	Sequence 10, Appli
27	363.5	22.5	527	PCT-US95-05008-10	Sequence 10, Appli

28	358.5	22.2	620	4	US-08-426-509A-9	Sequence 9, Appli
29	358.5	22.2	620	5	PCT-US95-05008-9	Sequence 9, Appli
30	356	22.0	625	1	US-08-391-615-3	Sequence 3, Appli
31	351.5	21.8	269	2	US-07-857-224B-79	Sequence 79, Appli
32	350	21.7	983	1	US-08-167-919A-10	Sequence 10, Appli
33	350	21.7	983	3	US-08-715-106-10	Sequence 10, Appli
34	349.5	21.6	304	2	US-08-701-191A-27	Sequence 27, Appli
35	349	21.6	983	2	US-08-449-645A-21	Sequence 21, Appli
36	349	21.6	983	2	US-08-702-367A-21	Sequence 21, Appli
37	349	21.6	983	5	PCT-US95-04681-21	Sequence 21, Appli
38	346.5	21.5	982	2	US-08-673-789-4	Sequence 4, Appli
39	345	21.4	276	2	US-07-857-224B-72	Sequence 72, Appli
40	344.5	21.3	983	1	US-08-162-809-16	Sequence 16, Appli
41	341.5	21.1	316	1	US-08-278-089A-16	Sequence 16, Appli
42	341.5	21.1	316	2	US-08-838-957A-15	Sequence 15, Appli
43	341.5	21.1	588	1	US-08-391-615-5	Sequence 5, Appli
44	339.5	21.0	787	4	US-09-188-930-334	Sequence 334, App
45	338.5	21.0	261	2	US-07-857-224B-58	Sequence 58, Appli

ALIGNMENTS

RESULT 1
US-09-529-279-4
; Sequence 4, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529,279
; CURRENT FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-279-4

Query Match	100.0%;	Score 1615;	DB 4;	Length 579;
Best Local Similarity	100.0%;	Pred. No. 5.2e-166;		
Matches 303;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MSTASASSSSSSSAGEMIEAPSOVLNFEIDYKEIEVEEWGVRGAFVYCKAKWRAKDV	60	
Db	1	MSTASASSSSSSSAGEMIEAPSOVLNFEIDYKEIEVEEWGVRGAFVYCKAKWRAKDV	60	
Qy	61	AIKQIESERKAFIVELRQLSRVNHENIVKLYGACLNPCVCLVMEYAEAGSLYNVLHGAE	120	
Db	61	AIKQIESERKAFIVELRQLSRVNHENIVKLYGACLNPCVCLVMEYAEAGSLYNVLHGAE	120	
Qy	121	PLPYVTAHAHMSWCLQCSQGVAYLHSMQPKALIHRLDKPPNLLLVAGTTLKICDFGTAC	180	
Db	121	PLPYVTAHAHMSWCLQCSQGVAYLHSMQPKALIHRLDKPPNLLLVAGTTLKICDFGTAC	180	
Qy	181	DIQTHMTNKGSAAMPAPEVPEGSNYSEKCDVFSWGIIILWEVITRRRKPDEIGGPAFRIM	240	
Db	181	DIQTHMTNKGSAAMPAPEVPEGSNYSEKCDVFSWGIIILWEVITRRRKPDEIGGPAFRIM	240	
Qy	241	WAVHNGTRPPLIKNLKPPIESLMTRCWKSPQSPSMEEIVKIMTHLMRYFPFGADEPLQY	300	
Db	241	WAVHNGTRPPLIKNLKPPIESLMTRCWKSPQSPSMEEIVKIMTHLMRYFPFGADEPLQY	300	
Qy	301	PCQ 303		


```
RESULT 5
US-09-221-527-5
; Sequence 5, Application US/09221527
; Patent No. 6146832
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,527
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-527-5

Query Match      26.8%; Score 433; DB 4; Length 455;
Best Local Similarity 36.7%; Pred. No. 2.3e-38;
Matches 99; Conservative 53; Mismatches 90; Indels 28; Gaps 11;

QY 27 NFEIDYKETEVEVGRGAFGVVCKAKW--RAKDVAIK---QIESESEKAFIVELROL 81
DB 7 SFVQIKFDDLOFFENCGGSGFSGVYRAKWISQDKEVAVKLLKIEKEAE-----IL 57

QY 82 SRVNHNPVVKLYGACLNLP--VCLVMEYAEAGGSLYNVLHG--AEPLPYTTAAHAMSWCLQC 137
DB 58 SVLSHRNIIQFYGVILEPPNYGIIVTEYASGLSDYDINSNRSEEM---DMDHIMTWATDV 114

QY 138 SOGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFTACDIQTHMTNNK--GSAAW 195
DB 115 AKGMHYLHMEAPVKVIHRLDKSRNVVIAADG-VLKICDFG-ASRPHNHTHMSLVGTFPW 172

QY 196 MAPEVFGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMW-AVHNGTRPPLIKN 254
DB 173 MAPEVIQSLPVSETCDTYSYGVVLWEMLTREVPFKGLEG--LQVAVLWVEKNERLTIPSS 230

QY 255 LPKPIESLMTRCWSKDPSPRSMEEIVKIM 284
DB 231 CPRSFAELLHQCEADAKKRPFSFKIISIL 260

; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-527-5

Query Match      26.8%; Score 433; DB 4; Length 455;
Best Local Similarity 36.7%; Pred. No. 2.3e-38;
Matches 99; Conservative 53; Mismatches 90; Indels 28; Gaps 11;

QY 27 NFEIDYKETEVEVGRGAFGVVCKAKW--RAKDVAIK---QIESESEKAFIVELROL 81
DB 7 SFVQIKFDDLOFFENCGGSGFSGVYRAKWISQDKEVAVKLLKIEKEAE-----IL 57

QY 82 SRVNHNPVVKLYGACLNLP--VCLVMEYAEAGGSLYNVLHG--AEPLPYTTAAHAMSWCLQC 137
DB 58 SVLSHRNIIQFYGVILEPPNYGIIVTEYASGLSDYDINSNRSEEM---DMDHIMTWATDV 114

QY 138 SOGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFTACDIQTHMTNNK--GSAAW 195
DB 115 AKGMHYLHMEAPVKVIHRLDKSRNVVIAADG-VLKICDFG-ASRPHNHTHMSLVGTFPW 172

QY 196 MAPEVFGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMW-AVHNGTRPPLIKN 254
DB 173 MAPEVIQSLPVSETCDTYSYGVVLWEMLTREVPFKGLEG--LQVAVLWVEKNERLTIPSS 230

QY 255 LPKPIESLMTRCWSKDPSPRSMEEIVKIM 284
DB 231 CPRSFAELLHQCEADAKKRPFSFKIISIL 260

; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-527-5

Query Match      26.8%; Score 433; DB 4; Length 455;
Best Local Similarity 36.7%; Pred. No. 2.3e-38;
Matches 99; Conservative 53; Mismatches 90; Indels 28; Gaps 11;

QY 27 NFEIDYKETEVEVGRGAFGVVCKAKW--RAKDVAIK---QIESESEKAFIVELROL 81
DB 7 SFVQIKFDDLOFFENCGGSGFSGVYRAKWISQDKEVAVKLLKIEKEAE-----IL 57

QY 82 SRVNHNPVVKLYGACLNLP--VCLVMEYAEAGGSLYNVLHG--AEPLPYTTAAHAMSWCLQC 137
DB 58 SVLSHRNIIQFYGVILEPPNYGIIVTEYASGLSDYDINSNRSEEM---DMDHIMTWATDV 114

QY 138 SOGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFTACDIQTHMTNNK--GSAAW 195
DB 115 AKGMHYLHMEAPVKVIHRLDKSRNVVIAADG-VLKICDFG-ASRPHNHTHMSLVGTFPW 172

QY 196 MAPEVFGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMW-AVHNGTRPPLIKN 254
DB 173 MAPEVIQSLPVSETCDTYSYGVVLWEMLTREVPFKGLEG--LQVAVLWVEKNERLTIPSS 230

QY 255 LPKPIESLMTRCWSKDPSPRSMEEIVKIM 284
DB 231 CPRSFAELLHQCEADAKKRPFSFKIISIL 260

; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-527-5

Query Match      26.8%; Score 433; DB 4; Length 455;
Best Local Similarity 36.7%; Pred. No. 2.3e-38;
Matches 99; Conservative 53; Mismatches 90; Indels 28; Gaps 11;

QY 27 NFEIDYKETEVEVGRGAFGVVCKAKW--RAKDVAIK---QIESESEKAFIVELROL 81
DB 7 SFVQIKFDDLOFFENCGGSGFSGVYRAKWISQDKEVAVKLLKIEKEAE-----IL 57
```

```
QY 82 SRVNHNPVVKLYGACLNLP--VCLVMEYAEAGGSLYNVLHG--AEPLPYTTAAHAMSWCLQC 137
DB 58 SVLSHRNIIQFYGVILEPPNYGIIVTEYASGLSDYDINSNRSEEM---DMDHIMTWATDV 114

QY 138 SOGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFTACDIQTHMTNNK--GSAAW 195
DB 115 AKGMHYLHMEAPVKVIHRLDKSRNVVIAADG-VLKICDFG-ASRPHNHTHMSLVGTFPW 172

QY 196 MAPEVFGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMW-AVHNGTRPPLIKN 254
DB 173 MAPEVIQSLPVSETCDTYSYGVVLWEMLTREVPFKGLEG--LQVAVLWVEKNERLTIPSS 230

QY 255 LPKPIESLMTRCWSKDPSPRSMEEIVKIM 284
DB 231 CPRSFAELLHQCEADAKKRPFSFKIISIL 260

; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-416-5

Query Match      26.8%; Score 433; DB 4; Length 455;
Best Local Similarity 36.7%; Pred. No. 2.3e-38;
Matches 99; Conservative 53; Mismatches 90; Indels 28; Gaps 11;

QY 27 NFEIDYKETEVEVGRGAFGVVCKAKW--RAKDVAIK---QIESESEKAFIVELROL 81
DB 7 SFVQIKFDDLOFFENCGGSGFSGVYRAKWISQDKEVAVKLLKIEKEAE-----IL 57

QY 82 SRVNHNPVVKLYGACLNLP--VCLVMEYAEAGGSLYNVLHG--AEPLPYTTAAHAMSWCLQC 137
DB 58 SVLSHRNIIQFYGVILEPPNYGIIVTEYASGLSDYDINSNRSEEM---DMDHIMTWATDV 114

QY 138 SOGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFTACDIQTHMTNNK--GSAAW 195
DB 115 AKGMHYLHMEAPVKVIHRLDKSRNVVIAADG-VLKICDFG-ASRPHNHTHMSLVGTFPW 172

QY 196 MAPEVFGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMW-AVHNGTRPPLIKN 254
DB 173 MAPEVIQSLPVSETCDTYSYGVVLWEMLTREVPFKGLEG--LQVAVLWVEKNERLTIPSS 230

QY 255 LPKPIESLMTRCWSKDPSPRSMEEIVKIM 284
DB 231 CPRSFAELLHQCEADAKKRPFSFKIISIL 260

; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-416-5

Query Match      26.8%; Score 433; DB 4; Length 455;
Best Local Similarity 36.7%; Pred. No. 2.3e-38;
Matches 99; Conservative 53; Mismatches 90; Indels 28; Gaps 11;

QY 27 NFEIDYKETEVEVGRGAFGVVCKAKW--RAKDVAIK---QIESESEKAFIVELROL 81
DB 7 SFVQIKFDDLOFFENCGGSGFSGVYRAKWISQDKEVAVKLLKIEKEAE-----IL 57

QY 82 SRVNHNPVVKLYGACLNLP--VCLVMEYAEAGGSLYNVLHG--AEPLPYTTAAHAMSWCLQC 137
DB 58 SVLSHRNIIQFYGVILEPPNYGIIVTEYASGLSDYDINSNRSEEM---DMDHIMTWATDV 114

QY 138 SOGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFTACDIQTHMTNNK--GSAAW 195
DB 115 AKGMHYLHMEAPVKVIHRLDKSRNVVIAADG-VLKICDFG-ASRPHNHTHMSLVGTFPW 172

QY 196 MAPEVFGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMW-AVHNGTRPPLIKN 254
DB 173 MAPEVIQSLPVSETCDTYSYGVVLWEMLTREVPFKGLEG--LQVAVLWVEKNERLTIPSS 230

QY 255 LPKPIESLMTRCWSKDPSPRSMEEIVKIM 284
DB 231 CPRSFAELLHQCEADAKKRPFSFKIISIL 260

; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-416-5

Query Match      26.8%; Score 433; DB 4; Length 455;
Best Local Similarity 36.7%; Pred. No. 2.3e-38;
Matches 99; Conservative 53; Mismatches 90; Indels 28; Gaps 11;

QY 27 NFEIDYKETEVEVGRGAFGVVCKAKW--RAKDVAIK---QIESESEKAFIVELROL 81
DB 7 SFVQIKFDDLOFFENCGGSGFSGVYRAKWISQDKEVAVKLLKIEKEAE-----IL 57

QY 82 SRVNHNPVVKLYGACLNLP--VCLVMEYAEAGGSLYNVLHG--AEPLPYTTAAHAMSWCLQC 137
DB 58 SVLSHRNIIQFYGVILEPPNYGIIVTEYASGLSDYDINSNRSEEM---DMDHIMTWATDV 114

QY 138 SOGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFTACDIQTHMTNNK--GSAAW 195
DB 115 AKGMHYLHMEAPVKVIHRLDKSRNVVIAADG-VLKICDFG-ASRPHNHTHMSLVGTFPW 172

QY 196 MAPEVFGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMW-AVHNGTRPPLIKN 254
DB 173 MAPEVIQSLPVSETCDTYSYGVVLWEMLTREVPFKGLEG--LQVAVLWVEKNERLTIPSS 230

QY 255 LPKPIESLMTRCWSKDPSPRSMEEIVKIM 284
DB 231 CPRSFAELLHQCEADAKKRPFSFKIISIL 260

; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-416-5

Query Match      26.8%; Score 433; DB 4; Length 455;
Best Local Similarity 36.7%; Pred. No. 2.3e-38;
Matches 99; Conservative 53; Mismatches 90; Indels 28; Gaps 11;

QY 27 NFEIDYKETEVEVGRGAFGVVCKAKW--RAKDVAIK---QIESESEKAFIVELROL 81
DB 7 SFVQIKFDDLOFFENCGGSGFSGVYRAKWISQDKEVAVKLLKIEKEAE-----IL 57
```

```

; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-245-5

```

Query Match	26.8%	Score 433	DB 4	Length 455
Best Local Similarity	36.7%	Pred. No. 2.3e-38		
Matches 99	Conservative 53	Mismatches 90	Indels 28	Gaps 11

```

Oy 27 NEEDIDYEIEIEEVEWVGAGVVCWAKM--RADVAIK--QIESESEKAFIVELROL 81
Db 7 SPVOIQKFDLDQFENFCGGGSGFSGYBAKMIQDKEVANKLLKIEKAE-----IL 57

Oy 82 SRVNHPIVKLYGACLN--VCLVMEVYEGSLYNVLHG--AEPLPYTAAHMSWLOC 137
Db 58 SVLSHRNIIQFAGVLLBPPNYGVLEVYASLSLSDVYINSNSEBM--DMDHITWATDV 114

Oy 138 SOGVYVHLSMOPKALIHRLKPNNLLVAGGTVIKCDFTACDIOHTMTNNK--GSAAM 195
Db 115 AKGMHYLHMEAPVKYIHHDLKSRNVIAADG-VLKIDCFG-ASAFPHNHTTHMSLVGFPM 172

Oy 196 MAPEVEGGSNVSEKDFVSGWIIIMEVITRRKPKDEIGGPAFRIMV-AVHNGTRPPLIKN 254
Db 173 MAPEVIOSLPVSSETDCTSYSGVILWEMETIREVPFKGLEG--LQVAMLVENKNERLTIPS 230

Oy 255 LPKPIESLMTRCWSDKPSDOPSMEEVIKIM 284
Db 231 CPRSAFELLHOCMEADAKKRPSFKOITIL 260

```

```

RESULT 9
US-09-163-115-5
; Sequence 5, Application US/09163115A
; Patent No. 6183962
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNT-050
; CURRENT APPLICATION NUMBER: US/09/163,115A
; CURRENT FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-163-115-5

```

Query Match	26.8%	Score 433	DB 43	Length 455
Best Local Similarity	36.7%	Pred. No. 2.3e-38		
Matches	99	Conservative 53	Mismatches 90	Indels 28
				Gaps 11
Qy	27	NFEEDIKYEIEVEEVVGCGAFGVCKAKW-RAKDVAIK--QIESSESRKAFIVELROL	81	
		: :		
		: :		
Db	7	SFVQIKFDLPDFFENCGGSGFSGVYAKKISDKEVAVKKLIEKEAE-----IL	57	
Qy	82	SRNNHNIVKLVGACLN--VCLVMYVYAGGSLYNNLHG--AEPLRYTAAHAMSWCLOG	137	
		: :		
		: :		
Db	58	SVLSHNNIIQFVILPEPNYGIVTETIASIGSLYDINSNRSEM--DMDHMTATGV	114	
Qy	138	SQGVAYLHSMQPKALIHRLDKPNNLLVAGGVTLKIDFQTACDIQHTMTNNK--GSAAM	195	
		: :		
		: :		
Db	115	AKGMHLHMEAPKVIHRDLKSRNVIIADG-VLKICDEG-ASRFNHTHTMSLVGTFFPM	172	
Qy	196	MAPEVEGGSYNSKCDVFSGGITLMEVIRRRKPFDEIGGAFRIM-AVHNGTRPLIN	254	
		: :		
		: :		
Db	173	MAPEVIGSLPVSTCTCTYSIGVLMEMLRVYFKELEG--LQVALVYVEKNERLIPSS	230	
Qy	255	LKPPIESLMTRCWSKDPQSPSPSMEELIVKIM	284	

```
Db 231 CPRSFAELHQCWEADAKKRPSFKQIISIL 260
```

```

RESULT 10
US-09-221-528-5
; Sequence 5, Application US/09221528
; Patent No. 6190874
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: WNI-050
; CURRENT APPLICATION NUMBER: US/09/1221,528
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-25
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-528-5

```

Query Match	26.8%	Score 433	DB 43	Length 455
Best Local Similarity	36.7%	Pred. No. 2,36-38		
Matches	99	Conservative	53	Mismatches 90; Indels 28; Gaps 11;
QY	27	NFEEDYKEIEVEEYVVGAGFVGVCKAW--RAKDVAIK--QISESEKAFIVELRQL	81	
DB	:	:	:	:
7	SVFOAIKFDLPDFOFNCGGSGFSGSVYRAKMIISODKEVAAKULKIKERKAE-----IL	57		
QY	82	SVNNPNIYKLYGACLN--VCLVMEVAGSGLYVNLHG--AEPPLYTAAMSMCLOC	137	
DB	:	:	:	:
58	SVLSHRNIIQFVGLVLEPPNYGVIVETASLSGLYDIYINSNSEEN---DMOHIMTADV	114		
QY	138	SGGVATISMQRKALIHRLKLPNNLLVAGGVYVKICDPGRACDIQTMINTNK--GSAAM	195	
DB	:	:	:	:
115	AGGMVHLHBAVKKIHRLLKSRNVIAADG-VLKICDFG--ASRPNNHTTHSLVGTPEW	172		
QY	196	MAPEVEGSGNSSEKDFVSMGCIILWEVITRRKPEDEIGAPFRIWM--AVHNGTRPPLIK	254	
DB	:	:	:	:
173	MAPEVIGSLPVSELCDDTSYGVVLWEMLTRVPPFGGLG--LQVAMLVVEKNERLTPSS	230		
QY	255	LKXPIESLMTRCWSKDPSPRSPMEIVIKM	284	
DB	:	:	:	:
231	CPRSFAELLHQCWEADAKKRPSPFKQIISIL	260		

```

RESULT 11
US-09-593-553-5
; Sequence 5, Application US/09593553
; Patent No. 6200770
; GENERAL INFORMATION:
; APPLICANT: Acteon, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNT-050
; CURRENT APPLICATION NUMBER: US/09/553,553
; CURRENT FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: 09/163,115
; PRIOR FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-593-553-5

Query Match      26.8%; Score 433; DB 4: Length 455;
Best Local Similarity 36.7%; Pred. No. 2,3e-38;
Matches 99; Conservative 53; Mismatches 90; Indels 28; Gaps 11

```

QY 27 NFEIDYKIEVEEVGRGAFVGVCKAKW--RAKDVAK--QIESESEKAFIVELROL 81
Db 7 SFVQIKPDDIQFENCNGGSGSVYRAKWSQDKEVAVKLLKIEKAE-----IL 57
QY 82 SRVNHNIKLYGACLN--VCLVMEYAGGSLYNVLHG--AEPLPYTAAHAMSWCLOC 137
Db 58 SVLSHRNIIOFYGVILEPPNYGIVTEVASLGSLYDINSNRSEEM---DMDHMTWATDV 114
QY 138 SQGVAYLHMQPKALIHRLDKPNNLLVAGGTVLKICDFGTACDIQTHMTNNK--GSAAW 195
Db 115 AKGMHYLHMEAPVKVHRLDKSRNVIAADG-VLKICDFG-ASRPHNHTHMSLVGTTPW 172
QY 196 MAPEVFGSNYSKCDVFSNGIILWEVITRRKPFDEIGGPAFRIMW-AVHNGTRPPLIKN 254
Db 173 MAPEVIQSLPVSCTDYSYGVVLWEMLTREVFPKLEG--LQVAVLWVEKNERLTIPSS 230
QY 255 LPKPISLMTRCWSKDPQSOPRSMEEIVKIM 284
Db 231 CPRSFAELLHQCEADAKRPSFKQIISIL 260

RESULT 12
US-09-221-237-5
; Sequence 5, Application US/09221237
; Patent No. 6214597
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,237
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-237-5

Query Match 26.8%; Score 433; DB 4; Length 455;
Best Local Similarity 36.7%; Pred. No. 2.3e-38;
Matches 99; Conservative 53; Mismatches 90; Indels 28; Gaps 11;
QY 27 NFEIDYKIEVEEVGRGAFVGVCKAKW--RAKDVAK--QIESESEKAFIVELROL 81
Db 7 SFVQIKPDDIQFENCNGGSGSVYRAKWSQDKEVAVKLLKIEKAE-----IL 57
QY 82 SRVNHNIKLYGACLN--VCLVMEYAGGSLYNVLHG--AEPLPYTAAHAMSWCLOC 137
Db 58 SVLSHRNIIOFYGVILEPPNYGIVTEVASLGSLYDINSNRSEEM---DMDHMTWATDV 114
QY 138 SQGVAYLHMQPKALIHRLDKPNNLLVAGGTVLKICDFGTACDIQTHMTNNK--GSAAW 195
Db 115 AKGMHYLHMEAPVKVHRLDKSRNVIAADG-VLKICDFG-ASRPHNHTHMSLVGTTPW 172
QY 196 MAPEVFGSNYSKCDVFSNGIILWEVITRRKPFDEIGGPAFRIMW-AVHNGTRPPLIKN 254
Db 173 MAPEVIQSLPVSCTDYSYGVVLWEMLTREVFPKLEG--LQVAVLWVEKNERLTIPSS 230
QY 255 LPKPISLMTRCWSKDPQSOPRSMEEIVKIM 284
Db 231 CPRSFAELLHQCEADAKRPSFKQIISIL 260

RESULT 13
US-09-291-839-2
; Sequence 2, Application US/09291839A
; Patent No. 6261818
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES THEREFOR

; FILE REFERENCE: MNI-068
; CURRENT APPLICATION NUMBER: US/09/291,839A
; CURRENT FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 835
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-291-839-2
Query Match 25.4%; Score 410; DB 4; Length 835;
Best Local Similarity 37.0%; Pred. No. 1.7e-35;
Matches 98; Conservative 50; Mismatches 97; Indels 20; Gaps 9;
QY 31 IDYKEIEVEEVGRGAFVGVCKAKWAKDVAIKQIE-----SESEKAFIVELROLSRVN 85
Db 458 LQLEIEFHEIIGSGSPGVKYGRCRNKIIVAKRYRANTVCSDVDVMFCREVSILCOLN 517
QY 86 HPNIVKLYGACLN--PVCLVMEYAGGSLYNVLHGAEPLPYTAAHAMSWCLOCQSGVA 142
Db 518 HPCVIQFVGACLNDPQSOFIVTQVISGSLFSLH--EQKRILDQLSKLIIAVDVAKGME 575
QY 143 YLHSM-QPKALIHRLDKPNNLLVAGGTVLKICDFGTACDIQ-----HMTNNKGSAAWMA 197
Db 576 YLHNLTP--IIHRDLNSHNILLYEDGHAV--VADFGESRFLQSLDEDNMTKQPGNLRMA 632
QY 198 PEVE-EGSNYSEKCDVFSNGIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNLP 256
Db 633 PEVTQCTRTYIKADVFSYALCLWEILUTGEIPFAHLKAPAAAADMAVHH--IRPPIGYSIP 691
QY 257 KPISLMTRCWSKDPQSOPRSMEEIV 281
Db 692 KPISLLIRGNACPEGRPESEVV 716
RESULT 14
US-08-205-018-2
; Sequence 2, Application US/08205018
; Patent No. 5554523
; GENERAL INFORMATION:
; APPLICANT: Reddy, Usharani R.
; APPLICANT: Pleasure, David
; TITLE OF INVENTION: No. 5554523el Protein Kinase, Nucleic Acid
; TITLE OF INVENTION: Sequences Encoding the Same and Methods Related
; TITLE OF INVENTION: Thereto
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and
; ADDRESS: No. 5554523ris
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/205,018
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Gaumond, Rebecca R.
; REGISTRATION NUMBER: 35,152
; REFERENCE/DOCKET NUMBER: CH-0488
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:

LENGTH: 668 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-205-018-2

Query Match 24.8%; Score 401; DB 1; Length 668;
 Best Local Similarity 33.5%; Pred. No. 1,2e-34;
 Matches 88; Conservative 54; Mismatches 99; Indels 22; Gaps 9;

QY 30 EIDYKEIEVEEYVGRGAFVVCCKAKWRADVAIKQIESSEKRAFIIVELRQLSRVNHNI 89
 DB 119 EVPEEILDLQWVGSAQAVFLGRHGEVAVKKYRDKE-----TDIKHLRKLGHPNI 173
 QY 90 VKLYGACLP--VCLVMEYAEAGSLYNNVLHGAEPLEPYTAHAMSGCLQCSQGVAYLHSM 147
 DB 174 ITPKGVCTQAPCCYCIIMEFCAQGLYEVLRAGRPV---TPSLVDMWMSGIAGMNYLHLH 230
 QY 148 QPKALIHRLDKPNNLLVAGTVLKCIDFGTACDIQTHMTNNK--GSAAMMAPEVFEESN 205
 DB 231 K--IIHRDLKSPN-MLITYDDVVKISDFGTSKELSDKSTKMSFAGTVAMMAPEVIRNEP 286
 QY 206 YSEKCDVFSMGITLMEVITRRKPFDEIGAPAFRIMAV-HNGTRPPLIKNLPKPIESLMT 264
 DB 287 VSEKVDIWSFGVVLWELLTGEIPIYKDVDSA--IIMGVGSNSLHLPVSSCPDGFKILR 344
 QY 265 RCWSKDPSQRPMSMEIYKIMTHL 287
 DB 345 QCWNSKPRNRPSPRQ---ILLHL 364

RESULT 15
 US-08-395-580-2
 ; Sequence 2, Application US/08395580
 ; Patent No. 5676945

GENERAL INFORMATION:

APPLICANT: Usharani R. Reddy, David Pleasure and the Children's
 TITLE OF INVENTION: Hospital of Philadelphia
 TITLE OF INVENTION: No. 5676945el Protein Kinase, Nucleic Acid
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and No. 5676945el
 STREET: One Liberty Place - 46th Floor
 CITY: Philadelphia
 STATE: PA
 COUNTRY: U.S.A.
 ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch disk, 720 KB
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Wordperfect 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/395,580
 FILING DATE: herewith
 CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/205,018
 FILING DATE: 01-MAR-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Rebecca L. Ralph (formerly Gaumond)
 REGISTRATION NUMBER: 35,152
 REFERENCE/DOCKET NUMBER: CH-0488
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 215-568-3100
 TELEFAX: 215-568-3439

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
 LENGTH: 859 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein

US-08-395-580-2

Query Match 24.8%; Score 401; DB 1; Length 859;
 Best Local Similarity 33.5%; Pred. No. 1,7e-34;
 Matches 88; Conservative 54; Mismatches 99; Indels 22; Gaps 9;

QY 30 EIDYKEIEVEEYVGRGAFVVCCKAKWRADVAIKQIESSEKRAFIIVELRQLSRVNHNI 89
 DB 119 EVPEEILDLQWVGSAQAVFLGRHGEVAVKKYRDKE-----TDIKHLRKLGHPNI 173
 QY 90 VKLYGACLP--VCLVMEYAEAGSLYNNVLHGAEPLEPYTAHAMSGCLQCSQGVAYLHSM 147
 DB 174 ITPKGVCTQAPCCYCIIMEFCAQGLYEVLRAGRPV---TPSLVDMWMSGIAGMNYLHLH 230
 QY 148 QPKALIHRLDKPNNLLVAGTVLKCIDFGTACDIQTHMTNNK--GSAAMMAPEVFEESN 205
 DB 231 K--IIHRDLKSPN-MLITYDDVVKISDFGTSKELSDKSTKMSFAGTVAMMAPEVIRNEP 286
 QY 206 YSEKCDVFSMGITLMEVITRRKPFDEIGAPAFRIMAV-HNGTRPPLIKNLPKPIESLMT 264
 DB 287 VSEKVDIWSFGVVLWELLTGEIPIYKDVDSA--IIMGVGSNSLHLPVSSCPDGFKILR 344
 QY 265 RCWSKDPSQRPMSMEIYKIMTHL 287
 DB 345 QCWNSKPRNRPSPRQ---ILLHL 364

Search completed: December 10, 2002, 03:52:38
 Job time : 42 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 10, 2002, 03:48:00 ; Search time 214 Seconds
(without alignments)
22.997 Million cell updates/sec

Title: US-09-830-144-2_COPY_1_303

Perfect score: 1615

Sequence: 1 MSTASAGSSSSSSAGEMIE.....MTHLMRYPPGADPLQVPCQ 303

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 103943 seqs, 16242309 residues

Total number of hits satisfying chosen parameters: 103943

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08 NEW PUB.pap.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT NEW PUB.pap.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06 NEW PUB.pap.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pap.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07 NEW PUB.pap.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pap.*
- 7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pap.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pap.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09 NEW PUB.pap.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pap.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10 NEW PUB.pap.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pap.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60 NEW PUB.pap.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1615	100.0	579	9	US-10-158-895-4
2	1615	100.0	590	9	US-10-158-895-15
3	433	26.8	455	10	US-09-757-982-5
4	432	26.7	1036	12	US-10-014-882-2
5	418.5	25.9	394	10	US-09-862-027-19
6	416	25.8	328	10	US-09-862-027-18
7	410	25.4	835	10	US-09-947-199-2
8	410	25.4	835	10	US-09-947-199-8
9	399	24.7	966	10	US-09-771-161A-197
10	393	24.3	263	10	US-09-840-704-5
11	392	24.3	850	10	US-09-904-389-2
12	369	22.8	425	10	US-09-828-313-29
13	363.5	22.5	527	10	US-09-977-269-10
14	360.5	22.3	277	10	US-09-882-166-4
15	359	22.2	278	9	US-09-842-582-4
16	359	22.2	278	10	US-09-797-039-13
17	359	22.2	278	10	US-09-922-138-18
18	359	22.2	278	10	US-09-922-138-27
19	359	22.2	278	10	US-09-910-150-17

20	359	22.2	278	10	US-09-910-150-31	Sequence 31, Appl
21	358.5	22.2	277	10	US-09-815-915-13	Sequence 13, Appl
22	358.5	22.2	620	10	US-09-977-269-9	Sequence 9, Appl
23	355	22.0	278	10	US-09-815-915-16	Sequence 16, Appl
24	351	21.7	273	10	US-09-922-138-11	Sequence 11, Appl
25	350	21.7	271	10	US-09-799-875-28	Sequence 28, Appl
26	350	21.7	272	10	US-09-780-949-4	Sequence 4, Appl
27	350	21.7	272	10	US-09-910-150-27	Sequence 27, Appl
28	349	21.6	983	10	US-09-771-161A-227	Sequence 23, Appl
29	348.5	21.6	279	10	US-09-799-875-23	Sequence 23, Appl
30	343	21.2	272	10	US-09-815-915-14	Sequence 14, Appl
31	333	20.6	507	10	US-09-977-269-2	Sequence 2, Appl
32	331.5	20.5	280	10	US-09-515-806-15	Sequence 15, Appl
33	329.5	20.4	928	8	US-08-578-684-2	Sequence 2, Appl
34	329	20.4	265	10	US-09-797-039-10	Sequence 10, Appl
35	329	20.4	675	9	US-10-186-399-3	Sequence 3, Appl
36	329	20.4	675	10	US-09-977-269-4	Sequence 4, Appl
37	329	20.4	764	10	US-09-923-302-714	Sequence 714, App
38	323.5	20.0	271	10	US-09-840-704-6	Sequence 6, Appl
39	320	19.8	1052	10	US-09-757-100B-2	Sequence 2, Appl
40	319.5	19.8	1276	10	US-09-982-610-24	Sequence 24, Appl
41	318.5	19.7	239	10	US-09-797-039-12	Sequence 12, Appl
42	318.5	19.7	245	10	US-09-815-915-15	Sequence 15, Appl
43	318.5	19.7	628	10	US-09-862-027-48	Sequence 48, Appl
44	318.5	19.7	1104	10	US-09-982-610-36	Sequence 36, Appl
45	317.5	19.7	212	10	US-09-834-496A-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-10-158-895-4
; Sequence 4, Application US/10158895
; Patent No. US20020155624A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-158-895-4

Query Match	100.0%;	Score	1615;	DB	9;	Length	579;
Best Local Similarity	100.0%;	Pred. No.	2.1e-124;				
Matches	303;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	MSTASAGSSSSSAGEMIEAPSOVLNFEEDYKEIEVEEVVGRGAFGVCKAKWRAKDV	60				
Db	1	MSTASAGSSSSSAGEMIEAPSOVLNFEEDYKEIEVEEVVGRGAFGVCKAKWRAKDV	60				
Qy	61	AIKQIESESRKAFIVELRQLSRVNHNPVVKLYGACLNPNVCLVMEYAEAGGSLYNVLHGAE	120				
Db	61	AIKQIESESRKAFIVELRQLSRVNHNPVVKLYGACLNPNVCLVMEYAEAGGSLYNVLHGAE	120				
Qy	121	PLPYTTAAHAMSWCLQCSQGVAYVLSHMQPKALIHRLDKPPNLLLVAGTGVLKICDFGTAC	180				
Db	121	PLPYTTAAHAMSWCLQCSQGVAYVLSHMQPKALIHRLDKPPNLLLVAGTGVLKICDFGTAC	180				


```

QY   77 -ELRQLSRVHNPIVKLYGACLN--VCLVMEYAEGGSLYNVLHGAEPLPYTA----- 127
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db   170 REARLFAMLRHPIIIEIRGVLCLOCPHLCLVLEFARGALNPAALAAANADPPRAPGPRA 229
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY   128 -----AHAM-SWCLOCQSGVAYLHSMQPALKIHRLDKPNLLLVAG-----GTVLKICD 175
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db   230 RRIPPHVLVNVAQTARGMLYLHEEAFVPIILHRDLKSNILLBKIEHDDICNKTLKITD 289
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY   176 FGTACD-IOTHMTNKGSAAMWAPVEVPEGSNSYSEKDVFSWGIIWEVITRRKPFEIGG 234
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db   290 FGLAREWHRTTKMSTAGTYAAMAEVIKSSLSFGSDINSYGVLINWELLTGEVDPYRGIDG 349
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY   235 PAPRIIMWAVHNGTRPPIIKNLPKPIESLMTRCWSKDPSQRSPMBEEIVKIMT 285
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db   350 LAVAYGAVNVKLTL-PIPSTCEPFPAKLMKECWQDPHIRSFALILEOLT 399
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

RESULT 5
US-09-862-027-19
; Sequence 19, Application US/09862027
; Patent No. US20020142428A1
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin R.
; TITLE OF INVENTION: NO. US20020142428A1el Kinases and Uses Thereof
; FILE REFERENCE: 35800/234862
; CURRENT APPLICATION NUMBER: US/09/862,027
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 09/345,473
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 394
; TYPE: EXT
; ORGANISM: Homo sapiens
US-09-862-027-19

Query Match          25.9%; Score 418.5; DB 10; Length 394;
Best Local Similarity 36.6%; Pred. No. 4.9e-27;
Matches 98; Conservative 47; Mismatches 100; Indels 23; Gaps 8;

QY   35 EIEVEEVGRGAGVGVCCKAKRAKDAIK--QIESESRKAFIGVELRQ----LSRVNHNP 88
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db   2 ELTLEEIIIGGGFGKYRAFWDIGDEVAVKAAARDHPDEDISQTIENVVRQEAKLFAMLKHPN 61
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY   89 IVKLYGACLN--VCLVMEYAEGGSLYNVLHGAEPLPYTHAAHMSWCLOCQSGVAYLHS 146
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db   62 IIALRGVCLKEFNLCVMEFAFGGPLNRVLSGKRIPPDI---LVNWAVQIARGWNLYLHD 117
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY   147 MQPKALIHRLDKPNLLLV-----AGGTVLKICDFGTACD-IOTHMTNKGSAAMWAP 198
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db   118 EAIVPIIIRDLKSNILLIQKVENGDSLKNKIUKITDFGLAREWHRTTKMSAAGTAWWAP 177
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY   199 EVFEGNSYSEKCDVFESWGIIWEVITRRKPFEIGGPAPRIIMWAV-HNGTRPPIIKNLPK 257
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db   178 EVIRASMFSGSDVMSYGVLLWELLTGEVFPFRGIDG--LRVAYGAMNKKALPIPSTCPE 235
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY   258 PIESLMTRCWSKDPSQRSPMBEEIVKIMT 285
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db   236 PPAKLMEDCWNPDPHSRPSFTNILDOLT 263
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

RESULT 6
US-09-862-027-18
; Sequence 18, Application US/09862027
; Patent No. US20020142428A1
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin R.
; TITLE OF INVENTION: NO. US20020142428A1el Kinases and Uses Thereof
; FILE REFERENCE: 35800/234862
; CURRENT APPLICATION NUMBER: US/09/862,027
; CURRENT FILING DATE: 2001-05-21

```

```
Db 576 YLHNLQTP--IHRDLNSHILLYEDGHAV-VADFESEFLOSLEDNDNTKOPGNLRMWA 632
Qy 198 PEVF-EGSNYSEKCDVFSWGIIIMEVITRRKPEDEIGAPFRIMAVHNGTRPPLIKNLP 256
Db 633 PEVFTCTRYTIKADVFSAALCLMEILTGEPFAHLKPAADAAVMAYHH-IRPPIGYSIP 691
Qy 257 KPESLMTRCMSKDPQSRPSMEIIV 281
Db 692 KPISILLIRGMWNACPGRPEPSEVV 716
```

RESULT 8

```
US-09-947-199-8
; Sequence 8, Application US/09947199
; Patent No. US20020127684A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CARC PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: MNI-068CP2
; CURRENT APPLICATION NUMBER: US/09/947,199
; PRIOR FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: 60/111,938
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/291,839
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 09/458,457
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 835
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-947-199-8
```

Query Match 25.4%; Score 410; DB 10; Length 835;

Best Local Similarity 37.0%; Pred. No. 5.7e-26; Matches 102; Conservative 50; Mismatches 100; Indels 24; Gaps 10;

```
Qy 20 EAPSOVLNEEIDYKEIEVEEVGARGAVGCAKAKRAKDAVIAKQIE-----SESERKAF 74
Db 451 ELPSRF---HQLSEIEPHEHIIIGSGFGKYKGRCKAKIYAKRYANTCSKSDVMF 506
Qy 75 IYELQSLSVNHPNIVKYGACL--NPYCLVMEYAEBSGLYVNHGAEPLPYTAAHAM 131
Db 507 CREVISILQCLNHPCVVQFVGAQLDPSQFAITQYISGSLFSLH--EOKRIIDLQSKL 564
Qy 132 SWCLQCSGVAYLHSM-OPKALIHRLDKPNNLLVAGGVLTICDPFGACDIQT---HM 186
Db 565 IIAVVAKAMEYLHSLTOP--IHRDLNSHILLYEDGHAV-VADFGESRFLQSLDEDNM 621
Qy 187 TNKKSAAMAEVVF-EGSNYSEKCDVFSWGIIIMEVITRRKPEDEIGAPFRIMAVHN 245
Db 622 TKQPGNLRWMAEVEVFTQCTRYTIKADVFSAALCLMEILTGEPFAHLKPAADAAVMAYHH 681
Qy 246 GTRPPLIKLKPESLMTRCMSKDPQSRPSMEIIV 281
Db 682 -IRPPIGYSIPKPISILLIRGMWNACPGRPEPSEVV 716
```

RESULT 9

```
US-09-771-161A-197
; Sequence 197, Application US/09771161A
; Patent No. US20020110811A1
; GENERAL INFORMATION:
; APPLICANT: Levine, et al.
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
; FILE REFERENCE: 802620-2005.1
; CURRENT APPLICATION NUMBER: US/09/771,161A
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/724,676
```

```
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 136776
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 135619
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 197
; LENGTH: 966
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-771-161A-197
```

Query Match 24.7%; Score 399; DB 10; Length 966;

Best Local Similarity 34.2%; Pred. No. 5.3e-25; Matches 90; Conservative 48; Mismatches 103; Indels 22; Gaps 9;

```
Qy 30 EIDYKEIEVEEVGARGAFVYCAKAKRAKDAVIAKQIESESEKAFIVELRQLSRVNHPI 89
Db 162 EVPEEISELQWLGSGAGAVFLGKFAEVAIKVREONE-----TDIKHLKLGHPNI 216
Qy 90 VKLYGACLN--VCLVMEYAEBSGLYVNHGAEPLPYTAAHAMSWCLQCSGVAYLHSM 147
Db 217 IAFGVTQAPCYCIIMEYCAHGOYLEVLRAKRT---TPRLIVDMSTGIASGMNYLHLH 273
Qy 148 QPKALIHRLDKPNNLLVAGGVLTICDPFGACDIQTHMTNKK--GSAAMAEVEFGSN 205
Db 274 K---IHRDLKSPN-VLVHTTDAVKISDFGTSKELSDKTSKSFAGTVAMMAPEVIRNEP 329
Qy 206 YSEKCDVFSWGIIIMEVITRRKPEDEIGAPFRIMAV-HNGTRPPLIKLKPESLMT 264
Db 330 VSEKVDIWSFGVVLWELTGELIPIKVDSSA--IIVGVSNSLHLVPSCTPDGFKILMK 387
Qy 265 RCWSKDPQSRPSMEIIVKIMTHL 287
Db 388 QTWQSKPRNRPSPRQ---TLMHL 407
```

RESULT 10

```
US-09-840-704-5
; Sequence 5, Application US/09840704
; Patent No. US20020122801A1
; GENERAL INFORMATION:
; APPLICANT: Dedhar, Shoukat
; TITLE OF INVENTION: Intergrin-Linked Kinase and its Uses
; FILE REFERENCE: KIN-2CON
; CURRENT APPLICATION NUMBER: US/09/840,704
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: 09/566,906
; PRIOR FILING DATE: 2000-05-09
; PRIOR APPLICATION NUMBER: US08/752,345
; PRIOR FILING DATE: 1996-11-19
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 263
; TYPE: PRT
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: Other
; LOCATION: (1)...(263)
US-09-840-704-5
```

Query Match 24.3%; Score 393; DB 10; Length 263;

Best Local Similarity 33.8%; Pred. No. 3.7e-25; Matches 90; Conservative 57; Mismatches 107; Indels 12; Gaps 8;

```
Qy 31 IDYKEIEVEEVGARGAFVYCAKAKRAKDAVIAKQIESESEKAFIVELRQLSRVNH 86
Db 1 IPWCDLINIKETIGAGSFVHRAEWHGSDVAVKIMEDQFAERVNEPLREVAIAIKRLRH 60
Qy 87 PNIVKLYGACLN--VCLVMEYAEBSGLYVNHGAEPLPYTAAHAMSWCLQCSGVAYL 144
```



```
Db 61 PNVLFMGAVTPQPNLSIVTEYLSRGLSYRLHLKSGAREQLDERRRLSMAYDVAKGNLYL 120
Qy 145 HSMQPKALIHRLDKPPNLLVAGGTVLKICDFG-TACDIQHTMNNK--GSAAMWAEVVF 201
Db 121 HNRNP-PIVHRDLKSPNLLVDKKYTV-KVCDPGLSRLKASTFLSSKSAAGTPEWMAPEVL 178
Qy 202 EGSNYSKCDVFSGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNLPKPIES 261
Db 179 RDEPSNEKSDVYSGFVILWELATLQOPGNL-NPAQVVAAGVFKCKRLLEIPRNLNPQAA 237
Qy 262 LMTRCWSKDPQSRPSMERIVKIMTHL 287
Db 238 IIEGCTWNEPWKRPSTFATIMDLLREL 263

RESULT 11
US-09-904-389-2
; Sequence 2, Application US/09904389
; Patent No. US20020129404A1
; GENERAL INFORMATION:
; APPLICANT: Clendennen, Stephanie K.
; APPLICANT: Schuster, Debra K.
; TITLE OF INVENTION: CTR1 HOMOLOGUE FROM MELON
; FILE REFERENCE: 4257-0029.30
; CURRENT APPLICATION NUMBER: US/09/904,389
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: US 60/218,307
; PRIOR FILING DATE: 2000-07-14
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 850
; TYPE: PRT
; ORGANISM: Cucumis melo
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (154)...(154)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-904-389-2

Query Match 24.3%; Score 392; DB 10; Length 850;
Best Local Similarity 32.1%; Pred. No. 1.7e-24;
Matches 93; Conservative 66; Mismatches 111; Indels 20; Gaps 12;

Qy 17 EMIEAPSOV-LNFE--IDYKEIEVEVVGCGAFGVCKAKWRAKDVAIKQIESES---E 70
Db 555 QLIRKPNELSLGLELIVPTDLDREKIGAGSGFTGYRGWHGSDVAVKILTEQDPHPE 614
Qy 71 R-KAFIVELRQLSRVNHPIVVKLYGACLP--VCLVMEYAEAGGSLYVNLH--GASPLPY 125
Db 615 RVNEFLUREVAIMKSLRHPNIVLFMGAVTPPNLSIVTEYLSRGLSYRLHLKSGVKDI--- 671
Qy 126 TAAHAMSWCLQCSQGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFG-TACDIQ 184
Db 672 DETFRINMAFDVAKGNLYLHRRDP-PIVHRDLKSPNLLVDKKYTV-KVCDPGLSRLKART 729
Qy 185 HMTNNK--GSAAMWAEVFEFGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWA 242
Db 730 FLSSKSAAGTPEWMAPEVLRLDEPSNEKSDVYSGFVILWELATLQOPWCNL-NPAQVVA 788
Qy 243 VHNGTTPPLIKNLPKPIESIMTRCWSKDPQSRPSMERIEIVKIMTHLMRYPP 292
Db 789 GFKGKRLDIPROVNPKNLASLIVACWADEPWKRPSTFSSIMETLKPMTKQAP 838
```

```
RESULT 12
US-09-828-313-29
; Sequence 29, Application US/09828313
; Patent No. US20020059662A1
; GENERAL INFORMATION:
; APPLICANT: COSTA e SILVA, OSWALDO DA
; APPLICANT: BOHNERT, HANS J.
```

```
; APPLICANT: THIELEN, NOCHA VAN
; APPLICANT: CHEN, ROUYING
; APPLICANT: SARRIA-MILLAN, RODRIGO
; TITLE OF INVENTION: PROTEIN KINASE STRESS-RELATED PROTEINS AND METHODS OF
; TITLE OF INVENTION: USE IN PLANTS
; FILE REFERENCE: 16313-0032
; CURRENT APPLICATION NUMBER: US/09/828,313
; CURRENT FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/196,001
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Physcomitrella patens
US-09-828-313-29
```

```
Query Match 22.8%; Score 369; DB 10; Length 425;
Best Local Similarity 33.8%; Pred. No. 5.8e-23;
Matches 94; Conservative 51; Mismatches 101; Indels 32; Gaps 10;

Qy 21 APSQVLNFEIDYKEIEVEVVGCGAFGVCKAKWRAKDVAIKQIESERKAFIV--- 76
Db 131 APSEI-----ELDTSELIGKGAFGIRKALMRGTPTVAVKTRPSSLNDRMVIKDFQ 181
Qy 77 -ELRQLSRVNHPIVVKLYGACL--NPVCLVMEYAEAGGSLYVNLHGAEPPLPYTTAAHMSW 133
Db 182 HEVQLLVKVRHPIVQFLGAVTRQRLMLVTEFLAGDLHQLLSRN---PNLAPDRIVK 238
Qy 134 CLQCSQGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFGTA--CDIQT-----HM 186
Db 239 ALDIARGMSYLHN-RSKPIIHRDLKPRN-IIVDEEHELKVGDFGLSKLIDVKLMDHVYKM 296
Qy 187 TNNKGSAAWMAPEVFEFGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNG 246
Db 297 TGGTGSYRYMAPEVFEHQPYDKSDVFSFGMILYEMFEGVAPPED--KDAYDAATLVARD 354
Qy 247 TRPPLIK--NLKPPIESLMTRCWSKDPQSRPSMERIEIVK 282
Db 355 DKRPENRAQTYPPQMKALIEDCWSPTYTPKRPPEIVK 392
```

```
RESULT 13
US-09-977-269-10
; Sequence 10, Application US/09977269
; Patent No. US20020082037A1
; GENERAL INFORMATION:
; APPLICANT: ULLRICH, AXEL
; APPLICANT: SURES, IRMINGARD
; APPLICANT: GISHIZKY, MIKHAIL
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
; FILE REFERENCE: 038602/1260
; CURRENT APPLICATION NUMBER: US/09/977,269
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 08/232,545
; PRIOR FILING DATE: 1994-04-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-977-269-10
```

```
Query Match 22.5%; Score 363.5; DB 10; Length 527;
Best Local Similarity 34.3%; Pred. No. 2.1e-22;
Matches 92; Conservative 48; Mismatches 111; Indels 17; Gaps 9;

Qy 30 EIDYKEIEVEVVGCGAFGVCKAKWRAK-DVAIKQI-ESESEKAFIVELRQLSRVNH 87
Db 260 EINPSELTFMRELGSGLFVVRLGKRWRAQYKVAIKAIREGAMCEEDFIEAKVMKLT 319
```


GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 10, 2002, 01:12:20 ; Search time 3085 Seconds
(without alignments)
2858.396 Million cell updates/sec

Title: US-09-830-144-2_COPY_1_303
Perfect score: 1615
Sequence: 1 MSTAASASSSSSSAGEMIE.....MTHLMRYFFGADEPLQYPCQ 303

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/cn2_1/USPTO spoal/US09830144/runat_04122002_141510_2911/app_query.fasta_1.455
-DB=GenEmbl -QFMT=fastap -SURFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09830144 @CN1.1.3637 @runat_04122002_141510_2911 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WATN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DLEXT=7

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.on.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1615	100.0	1704	6	E38399 NF-Kappa B
2	1615	100.0	1705	9	AB009358 Homo sapi
3	1615	100.0	1745	9	AF218074 Homo sapi
4	1615	100.0	2757	9	BC017715 Homo sapi
5	1615	100.0	2769	6	AX377912 Sequence
6	1615	100.0	2769	9	AB009356 Homo sapi
7	1615	100.0	2785	6	E38397 NF-kappa B
8	1615	100.0	2850	9	AB009357 Homo sapi
9	1615	100.0	2866	6	E38398 NF-kappa B
10	1609	99.6	2443	10	MUSTAK1
11	1609	99.6	3107	10	BC006665 Mus muscu
12	1490.5	92.3	2812	5	U92030 Xenopus lae
13	1414	87.6	135147	2	AC114407 Mus muscu
14	818	50.7	3349	3	AF199466 Drosophil
15	818	50.7	3386	3	AY051953 Drosophil
16	689.5	42.7	32564	2	AC014558 Drosophil
17	689.5	42.7	164942	3	AC011758 Drosophil
C 18	689.5	42.7	302303	3	AE003571 Drosophil
C 19	482.5	29.9	3138	9	HSMSTM
20	482.5	29.9	3454	6	AX337846 Sequence
21	482.5	29.9	3454	9	HARNMLK2
22	454	28.1	3931	9	AF251442 Homo sapi
23	442	27.4	2120	6	AR119790 Sequence
24	442	27.4	2120	6	AR126750 Sequence
25	442	27.4	2120	6	AR128910 Sequence
26	442	27.4	2120	6	AR130841 Sequence
27	442	27.4	2120	6	AR138886 Sequence
28	442	27.4	2120	6	AR141354 Sequence
29	437	27.1	8125	3	AF481923 Dictyoste
30	435	26.9	1429	10	AB049732 Mus muscu
31	435	26.9	3146	10	AB049731 Mus muscu
32	434.5	26.9	1267	8	AY133876 Arabidops
33	434.5	26.9	1798	8	AY046026 Arabidops
34	434	26.9	1370	9	AF325454 Homo sapi
35	434	26.9	1571	9	BC001401 Homo sapi
36	434	26.9	1631	8	AY085535 Arabidops
37	434	26.9	1795	8	AY125513 Arabidops
38	434	26.9	2251	9	AK056310 Homo sapi
39	434	26.9	2631	9	AB030034 Homo sapi
40	434	26.9	3767	9	AF251441 Homo sapi
41	434	26.9	3884	9	AF480461 Homo sapi
42	434	26.9	7195	9	AF480462 Homo sapi
43	433	26.8	1365	6	AR119791 Sequence
44	433	26.8	1365	6	AR126751 Sequence
45	433	26.8	1365	6	AR128911 Sequence

ALIGNMENTS

RESULT 1

E38399	E38399	1704 bp	DNA	linear	PAT 31-JAN-2002
LOCUS	NP-kappa B activation inhibitory drug targeting TAK1 and method for				
DEFINITION	identifying the same.				
ACCESSION	E38399.1	GI:18626979			
VERSION	E38399.1	JP 2000197500-A/5.			
KEYWORDS	unidentified.				
SOURCE	unidentified.				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 1704)				
AUTHORS	Sugita,N., Sakurai,H., Kageyama,N. and Haegawa,H.				
TITLE	NP-kappa B activation inhibitory drug targeting TAK1 and method for				
JOURNAL	identifying the same				
	Patent : JP 2000197500-A 5 18-JUL-2000;				
	TANABE SEIRYAKU CO LTD				
COMMENT	OS Unidentified				
	PN JP 2000197500-A/5				
	PD 18-JUL-2000				
	PF 04-FEB-1999 JP 1999026803				
	PR				
	PI NAOHISA SUGITA,HIROAKI SAKURAI,NORIKO KAGEYAMA, PI HIROSHI				
	HASEGAWA				
	PC C12Q1/48,A61K31/00,A61K31/00,A61K45/00,C12N5/10,C12N9/99, PC				
	C12Q1/02,				
	PC G01N33/15,G01N33/50,G01N33/566//C12N15/09,C12Q1/68,(C12N15/09,				
	PC C12R1:91),				
	PC C12N5/00,C12N15/00,(C12N15/00,C12R1:91)				
	CC Strandedness: Double;				
	CC Topology: Linear;				
	FM Key	Location/Qualifiers			
	FT source	1..1704			
		/organism='Unidentified'.			
FEATURES	source	1..1704			
		/organism="unidentified"			
		/db_xref="taxon:32644"			
BASE COUNT	511 a 381 c 401 g 411 t				
ORIGIN					
Alignment Scores:					
Pred. No.:	2,52e-149	Length:	1704		
Score:	1615.00	Matches:	303		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Best Query Match:	100.00%	Indels:	0		
DB:	6	Gaps:	0		
US-09-830-144-2_COPY_1_303 (1-303) x E38399 (1-1704)					
QY	1 MetSerThrAlaSerAlaIaIaSerSerSerSerSerAlaGluGluMetIleGlu 20				
DB	1 AGGTCTACAGCCCTGCCCCCTCTCTCTCTCTCTCTCTGCGCCGTGAGATCGAA 60				
QY	21 AlaProSerGluValIleuasnPhegiugluIleasPtyrIleGluIleGluValGluGlu 40				
DB	61 GCCCTTCCCAAGTCTCTCAACTTGAAGAAGATGACTTACAAGAGATCGAGGTGAAGAG 120				
QY	41 ValValGlyArgGlyValAlaPheGlyValValCysGlyAlaIaIaSTPrArgAlaIysAspVal 60				
DB	121 GTTGTGTGAAGAAGGAGACCTTTGGAGTGTGTTGCGMAAGCTTAAGTGAAGACCAAAAGATGTT 180				
QY	61 AlaIleIysGluIleIleGluSerGluSerGluArgLysAlaAlaPheIleValGluIleuAArgGln 80				
DB	181 GCTATTAAACAAATATGAAGATGATCTGAGAGAGAAAGCGTTTATTGTAGAGCTTCGGCAG 240				
QY	81 LeuSerArgValAsnHisIspProsnIleValIysLeuIleTyrValaCysLeuAsnProval 100				
DB	241 TTATCCCGTGTGAACCATCTTAATATTGTGAAGCTTTATGAGACCTGGAATCCAGTG 300				
QY	101 CysLeuValMetGluTyrTralagIugIyGlySerIeTyrAsnValIleuHisGlyAlaGlu 120				
DB	301 TGTCTGTATGAATAATGCTGAAGGGGCTCTTATATPATATGTGCGATGGTGTCTGAA 360				

QY	121	Proleuprotin	ThyThAlaAlaHisAlaMetSerTrpCysLeuGlnCysSerGlnGly	140
Db	361	CCATTGCCATTATTAACTGCTGCCACGCAATGATGGTGGTTTACAGTGTCCCAAGGA	420	
QY	141	ValAlaTyrLeuHisSerMetGlnProValAlaLeuLeuLeuHisArgAspLeuysPropio	160	
Db	421	GGGCGTTATCTTCACAGATCAACCCAAAGCGTTATTCACGGGACCTGAAACACCA	480	
QY	161	AsnLeuLeuLeuValAlaGlyGlyThrValLeuysIleCysAspPheGlyThraIacys	180	
Db	481	AACTACCTGCTGGTTCAGAGGGGAGCAGTCTTAAAAATTGTATTTGTGTACAGCTGT	540	
QY	181	AspIleGlnThHisMetThrAsnAsnysGlySerAlaAlaTrpMetAlaProGluVal	200	
Db	541	GACATTCACACACACATATACCAATPACAGGGGAGTGGCTTGATGACCTGAAGTT	600	
QY	201	PheGlnGlySerAsnTyrSerGlnIleCysCysAspValPheSerTrpAlaIleIleLeuTrp	220	
Db	601	TTTGAAGGATAGTAAATTACACGTGAAAATATGACCTCTTACGCTGGGGTATATCTTTGG	660	
QY	221	GluValIleThrArgArgIysProPheAspGluIleGlyGlyProAlaPheArgIleMet	240	
Db	661	GAGTGATTAACCGCTGCGAAGAACCTTTGATGATGATGGTGGCCACCTTCCGAATCAAG	720	
QY	241	ThrAlaValHisAsnGlyThrArgProProleuIleLysAsnLeuProLysProIleGlu	260	
Db	721	TGGGCTGTTCAATTAAGTACTGACACACATGATAAAAATTTACTTAAGCCCATTTGAG	780	
QY	261	SerLeuMetThrArgCysTrpSerLysAspProSerGlnArgProSerMetGluGluIle	280	
Db	781	AGCTGATAGCTGGTTGGTCTAAAGATCTTCCACAGGCCCTTCAATGAGGAAATT	840	
QY	281	ValIysIleMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyr	300	
Db	841	GTGAAATATATATACCTCATTTGGGGTAACTTTCAGAGCAGATGAGCCATTACAGTAT	900	
QY	301	ProCysGln 303		
Db	901	CCTTGTCTAG 909		
RESULT 2				
AB009358		1705 bp	mRNA	linear
LOCUS				PRI 20-JAN-2001
DEFINITION			Human sapiens mRNA for TGF-beta activated kinase 1c, complete cds.	
ACCESSION			AB009358	
VERSION			AB009358.2	GI:8978251
KEYWORDS			TAK1c; TGF-beta activated kinase 1c.	
SOURCE			Human sapiens cell_line:HeLa cDNA to mRNA, clone: pT7TAK1c.	
ORGANISM			Human sapiens	
REFERENCE			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE			1 (sites)	
AUTHORS			Sakurai,H., Shigemori,N., Hasegawa,K. and Sugita,T.	
TITLE			TGF-beta-activated kinase 1 stimulates NF-kappa B activation by an NF-kappa B-inducing kinase-independent mechanism	
JOURNAL			Biochem. Biophys. Res. Commun. 243 (2), 545-549 (1998)	
MEDLINE			98153801	
REFERENCE			2 (sites)	
AUTHORS			Dempsey,C.E., Sakurai,H., Sugita,T. and Guesdon,F.	
TITLE			Alternative splicing and gene structure of the transforming growth factor beta-activated kinase 1	
JOURNAL			Biochim. Biophys. Acta 1517 (1), 46-52 (2000)	
MEDLINE			20568765	
REFERENCE			3 (bases 1 to 1705)	
AUTHORS			Sakurai,H., Shigemori,N., Hasegawa,K., Sugita,T. and Guesdon,F.	
TITLE			Direct Submission	
JOURNAL			Submitted (01-DEC-1997) Francois Guesdon, University of Sheffield, Royal Hallamshire Hospital, Division of Molecular and Genomic Medicine, Functional Genomics Group, Glossop road, Sheffield S10 2JF, United Kingdom (E-mail:f.a.guesdon@sheffield.ac.uk, Fax:+44-114-271-3846)	
COMMENT			On Jul 8, 2000 this sequence version replaced gi:2924627.	

FEATURES	Location/Qualifiers		Db	
	source			
CDS	1..1705		541	GACATTCAGACACATGACCAATAACAAGGGAGTGCTGCTGGATGGGCACCTGAAGTT
	/organism="Homo sapiens"		201	PheGluGlySerAsnTyrSerGluLysCysAspValPheSerTyrGlyValIleLeuTyr
	/db_xref="taxon:9606"			
	/clone="pT7TAK1c"		601	TTTGAAGTAGTAAATTACAGTGAATAATGTCAGCTCTTCAGCTGGGTATTATTCTTTGG
	/cell_line="HeLa"			
	1..1557		221	GluValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMet
	/notes="TAK1c"		661	GAAGTGATAACGCGTCGGAAACCCCTTTGATGAGATTGGTGGCCAGCTTTCGGAATCATG
	/codon_start=1			
	/product="TGF-beta activated kinase 1c"		241	TyrAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGlu
	/protein_id="BAA25027.2"		721	TGGGCTGTTCATAATGGTACTCGACCACTGATAAAAAATTTACCTAAGCCCATTTGAG
	/db_xref="GI:8978252"			
	/translations="MSTASAASSSSSSAGEMIEAPSQVLNFEEDYKEIEVEEVVGR		261	SerLeuMetThrArgCysTyrSerLysAspProSerGlnArgProSerMetGluGluIle
	GAGVYCKAKRWAKDVAIKQIIESESEKAFIVELRQLSRVNHPIVKLYGACLNVPCL			
	VMEYARGSLYNVLHGAEPLPYTAAHAMSWCLQCSQGVAYLHSMQPKALIHRLDKPP		781	AGCCTGATGACTCGTTGTTGTTCTAAAGATCCTTCCAGCGCCCTTCAATGGAGAAATT
	NLLVAGGTVLKICDFGTACDIOTHTNNKGAAMWAPVEFGSNYSEKCDVFSWGI			
	LWEVITRRKPFDEIGGAPRIMWAVHNGTRPPLIKNLPKPIESLAWRCWSKDPSPRS		281	VallysIleMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyr
	MEEVIKIMTHLMRYFPGADEPLQYPCOYSDGQSNASTSGFMIDIASTNNYNSKDTN			
	MEQVPAINDTIKKLESKLKNQAQKQSESLGASRGSSVESLPPTSEBGRKMSADM		301	ProCysGln
	SETEARIAATTAYSKPKGRHRTASFENILDVPEIVISGNQPRRRSIQDLTVTTEP			
	GOVSSRSRSPSVMITTSPTSEKPTRSHPWTPDDSDTDTNGSDNSIPMAYLTLDHQLQ		841	GTGAAAAATAATGACTCACTTCATGCGGTACTTTCAGGAGCAGATGAGCCATTACAGTAT
	ARTSCRTGPG"			
BASE COUNT	511 a 381 c 402 g 411 t		301	ProCysGln 303
ORIGIN			901	CCTTGTGACG 909
Alignment Scores:				
Pred. No.:	2,52e-149	Length:	1705	
Score:	1615.00	Matches:	303	
Percent Similarity:	100.00%	Conservative:	0	
Best Local Similarity:	100.00%	Mismatches:	0	
Query Match:	100.00%	Indels:	0	
DB:	9	Gaps:	0	
US-09-830-144-2_COPY_1_303 (1-303) x AB009358 (1-1705)				
Qy	1	MetSerThrAlaSerAlaAlaSerSerSerSerSerSerAlaGlyGluMetIleGlu	20	
Db	1	ATGFTACAGCCTCGCGCGCTCCTCCTCCTCGCTTCGCGCGGTGAGATGATCGAA	60	
Qy	21	AlaProSerGlnValLeuAsnPheGluGluIleAspTyrLysGluIleGluValGluGlu	40	
Db	61	GCCCCCTCCAGGTCCTCACTTTGAAGATCGACTACAGGAGATCGAGGTGGAAGAG	120	
Qy	41	ValValGlyArgGlyAlaPheGlyValValCysLysAlaLysTyrArgAlaLysAspVal	60	
Db	121	GTTCGTGGAAGAGAGCCTTTGAGATTGTTTGCAAACTAAGTGGAGAGCAAAAGATGT	180	
Qy	61	AlaIleLysGlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln	80	
Db	181	GCTATTAAACAAATAGAAAGTGAATCTGAGAGGAAGCGTTTATTGTAGAGCTTCGGCAG	240	
Qy	81	LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnProVal	100	
Db	241	TTATCCCGTGTGAACCATCTTAATATTGTAAGCTTTATCGAGCGCTTGAATCCAGTG	300	
Qy	101	CysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGlyAlaGlu	120	
Db	301	TGTCCTGTGATGATATGCTGAAGGGGCTCTTATATATATGTGTCATGTGCTGAA	360	
Qy	121	ProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeuGlnCysSerGlnGly	140	
Db	361	CCATTGCCATATTATACGTGCTGCCACGAATGAGTTGGTGTATTACAGTGTCCCAAGGA	420	
Qy	141	ValAlaTyrIleuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysProPro	160	
Db	421	GTGCGCTTATCTTACAGCATGCAACCCAAACGCGCTAATTCACAGGACCTGAAACACCA	480	
Qy	161	AsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCys	180	
Db	481	AACTTACTGTGTTGAGGGGGACAGTTCTAAAAATTTGTAATTTGGTACAGCGCTGT	540	
Qy	181	AspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyrMetAlaProGluVal	200	

[illegible]

RESULT 4
 Db 907 CTTGTCTGAG 915
 LOCUS BC017715
 DEFINITION Homo sapiens, mitogen-activated protein kinase kinase kinase 7,
 ACCESSION BC017715
 VERSION BC017715.1 GI:17389342
 KEYWORDS MGC.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 2757)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (03-DEC-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 COMMENT Contact: MGC help desk
 Email: gcgabs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-sngc.stanford.edu>
 Contact: (Dickson, Mark) mcd@paxill.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>
 Series: IRAC Plate: 22 Row: 1 Column: 18
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gt: 4507360.
 FEATURES
 source location/Qualifiers
 1..2757
 /organism="Homo sapiens"
 /db_xref="LOCUSID:6885"
 /db_xref="taxon:9606"
 /clone="MGC:21263 IMAGE:3906837"
 /tissue_type="Uterus, leiomyosarcoma"
 /clone_id="NH_MGC_71"
 /lab_host="DH10B"
 /note="vector: pCMV-Sport6"
 198..1937
 /codon_start=1
 /product="mitogen-activated protein kinase kinase kinase
 7"
 /protein_id="AAH17715.1"
 /db_xref="GI:17389343"
 /translation="MSPSASASSSSSSSAGEMIEAPQVNLFEIDYKEIEVEVGR
 GARVVCARKAKRAKDVAIKOIESESEKATIVELRQLSRNVNPIVLYGACLNPCCL
 VMEAEBSGLYNVLHGAEPRPYTAAMNSMCCQSGVAYLHSMOKALIHRLKPR
 MLLNAGGIVLKICDFETACDICIHTMTNKGSAAMAPVEYBGSNYSKCVFSGWGIIL
 LMWEYTRKRPDEIGGAPAFIRMAVNHGTPPLIKLPRKIESLMTCSWSDPQRPSP
 MEELYKIMTLMREYFGADPPLQVPCYSEGGNSATSTGSFMDISTNSKSDNTN
 MEQVPAINTDTIKRLSKLNKQAQSESRSLGASGASVEELPTSGSKMSADM
 SELEARIATNTGGQPRRSIDQLTVTGCTPGGVSSSSPSVMITITPTSPKPTR
 SHPTPDPDSTDITNGSDNSIMAYITLTDHQLQPLAPCNSESNAAVEFQCKMADEYIKK
 VQETIALLRQKQELVAELDDQDEKDOQNTSLRVLQEHKKLLDENKSLSTYQQCKQKLE
 VIRSQQKRGQTS"
 BASE COUNT 819 a 566 c 647 g 725 t
 ALIGNMENT Scores: 4.46e-149 Length: 2757
 Pred. No.:

Db	763	TTTGAAGGTAGAATTAACGATCAAAAAATGTGACGCTTCAGCGTGGGGTATTATCTTTGG	822
Oy	221	GTVALLIIEThArGrRgLySPropheaspguileglyGVlProblabheargllemeT	240
Db	823	GAAGCATTAACGGCTGGAAAACCCCTTTGATRGAAATTGGTGGCCAACCTTCCGAATCATG	882
Oy	241	TTPAlavahIsasnlylThrArgProPoleuilelYsAsnleuProLysprolleiglu	260
Db	883	TGGCGCTGCATTAATGTACTGCACCACCATGATATAAATAATTACTTAAGCCCATTTGAG	942
Oy	261	SerleuMeThArGcStrtPserLySaapProSerGlharGProSerMetglugluile	280
Db	943	AACCGATACACGCTGTGTGGCTTAAGAATCCTTCCCACGCCCTTCAATGGAGGAAATT	1002
Oy	281	VallYslleMeThrHisleuMeArGTrYPheProGLVAlASpdluProleughIntyr	300
Db	1003	GTCAAAATATATACCTCACCTTGATCCGGTACTTTCACAGSGACGAATGAGCCATTACAGTAT	1062
Oy	301	ProCySgin 303 	
Db	1063	CCTGTGCAG 1071	
RESULT 6			
LOCUS	AB009356		
DEFINITION	AB009356	2769 bp mRNA linear PRI 04-MAR-1998	
ACCESSION	AB009356	Homo sapiens mRNA for TGF-beta activated kinase 1a, complete cds.	
KEYWORDS	AB009356.1	GI:2924623	
SOURCE		TAKIYA, TGF-beta activated kinase 1a. Homo sapiens lung cDNA to mRNA, clone_1lb:lambda gt11	

ORGANISM	Homo sapiens
REFERENCE	Eukaryota, Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (sites)
AUTHORS	Sakurai, H., Shigemori, N., Hasegawa, K. and Sugita, T.
TITLE	TGF-beta-activated kinase 1 stimulates NF-kappa B activation by an NF-kappa B-inducing kinase-independent mechanism
JOURNAL	Biochem. Res. Commun. 243 (2), 545-549 (1998)
MEDLINE	98153801
REFERENCE	2 (bases 1 to 2769)
AUTHORS	Sakurai, H.
TITLE	Direct Submission
JOURNAL	Submitted (01-DEC-1997) Hiroaki Sakurai, Lead Generation research laboratory, Tanabe Seiyaku Co. Ltd., 16-89, Koshima-3-chome, Yodogawa-ku, Osaka, Osaka 532, Japan (E-mail:hsakurai@tanabe.co.jp, Tel:++81 6 300 2571, Fax:++81 6 300 2593)

FEATURES	Location/Qualifiers
source	1. .2769
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone="PBSTAK1a"
	/tissue_type="lung"
	/clone_lib="Lambda gt11"
CDS	163. .1902

```

/note="TAKIA"
/codon_start=1
/product="NGF-beta activated kinase 1a"
/protein_id="BAA25025.1"
/db_xref="GI:2924624"
/translation="MSTASASASSSSSAGSEMIBAPQVINFEEIDYKETEIVEEVGR
GAFVNSCAKAKRMADVAIKQIESERKKAFLVELRQSRVNHPIVLKYGACINPCL
VMEVAGSGSLVNLHGAEPLPYTPAALAMNSCGLQVAVLHMQKALIHRLKRL
NLLVAGGTVALIKCDPFGATCIQTHMNNKSGALMMAPEVNGEVSREKDDPWSGII
LMVEITRRKPRDEIGCAPRIMAAVNHGTRPLIKULPKPIESLMTCSKDSORS
MEQVATNHTLMRYFPGADEPQVPCQVSDGSGNSATSTGSEVDAISATNSKSDTN
SEIDARIATGNGQPRRRSIQDLTVGTSYSGSSSSSPVSMITTSPTSEKPR
SHETPPDSTDINGDSNSIPAAVLTLLHOLQPLAPCNSKESMAVFQHKCMQEVKK
VQTEIALLRKQELVAELDDQDEKQDQNTSLRYEHKKLIDENKSLSTYYQCKKQLE
VIRSQQCKRQGST"
BASE COUNT      811 a      565 c      640 g      753 t

```

ORIGIN	
Alignment Scores:	
Pred. No.:	4,48e-149
Score:	1615.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	9
Length:	2769
Matches:	303
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-830-144-2_COPY_1_303 (1-303) x AB009356 (1-2769)
QY 1 MetSerThrAlaSerAlaAlaSerSerSerSerSerAlaGlyValMetIleGlu 20
Dp 163 ATGTCTACAGCCTTGGCGGCTCTCCCTCTCTCTCTCTCTGGCGGTAGATGATCGAA 222
QY 21 AlaProSerGlnValLeuAsnPheGluGlnIleAspTyLysGluIleGluValGluGlu 40
Dp 223 GCCCCTTCACAGGCTCTCAACTTTGAAGAAGATCGATTACAAGAGATCGAGGTGAAAG 282
QY 41 ValValGlyArgGlyAlaPheGlyValValCysLysAlaLysTrpArgAlaLysAspVal 60
Dp 283 GTTGTTGGAAGAGAGAGCCTTTGAGATTGTTGGCAAAGTAAGTGAAGCAAAAGATGTT 342
QY 61 AlalLeuSerGlnIleGluSerGluSerGluValGlyValPheIleValGlyLeuArgGln 80
Dp 343 GCTATTMAAATAATAGAAAGTGAATCTGAGAGAAAGCCTTATTGTATGACCTGGCGAG 402
QY 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyGlyAlaCysLeuAsnProVal 100
Dp 403 TTATCCCGTGAACCATCTTAATATTGTAAAGCTTTATGAGACCTGGCTGATCAAGTG 462
QY 101 CysLeuValMetGluTyraIaGluGlyGlySerLeuTyraSnValLeuHisGlyAlaGlu 120
Dp 463 TGCTTGATGAGAAATATGCTGAAAGGGGCGCTTTATATATATGCTGATGCTGTGA 522
QY 121 ProLeuProLyrrTyrrThraAlaAlaHisAlaMetSerTrpCysLeuGlnCysSerGlnTy 140
Dp 523 CCATTGCCAATTAATATCTGCTGCCACGCMAATGAGTGTGTTTACAGTGTCCCAAGGA 582
QY 141 ValAlaTyrrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysProPro 160
Dp 583 GTGGCTTATTTTACAGCATGCAACCMAAGGCTTAATTCACAGGACCTGAAACACCA 642
QY 161 AsnLeuLeuLeuValaIaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCys 180
Dp 643 AACTTAGCTGGTTCACAGGGGAGACAGTCTTAAAAATTGTGATTTTGGTACACCTGT 702
QY 181 AspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGluVal 200
Dp 703 GACATTGACACACATGACCAATTAACAAGGGAGTGTGCTTGATGGCACCTGAAGTT 762
QY 201 PheGluGlySerAsnTyrrSerGlnLysCysAspValPheSerTrpGlyTyrIleLeuTrp 220
Dp 763 TTTGAAGGATGATTAATTAACGTAAGAAAATGTGACGCTTCAGCTGGGGATTAATCTTTGG 822
QY 221 GluValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMet 240
Dp 823 GAAGTATACCGGCTGGAAACCTTTGATGATGGATGGGCCCAAGCTTTCCGAATCATG 882
QY 241 TrpAlaValHisAsnGlyThrArgProLeuIleLysAsnLeuProLysProIleGlu 260
Dp 883 TGGGCTGTTCATATGATGCTACTGCACACCATGATAAAAAATTATTAACTAAGCCATTAG 942
QY 261 SerLeuMetThrArgCysTrpSerLysAspProSerGlnArgProSerMetGluGluIle 280
Dp 943 AGCCTATATACCTCGTTGTGTGCTTAAAGTCTCTTCCAGCGCCCTTCATAGAGCAAAATT 1002
QY 281 ValLysIleMetThrHisLeuMetArgTyrrPheProGlyAlaAspGluProLeuGlnTyrr 300
Dp 1003 GTGAAATTAATGATCTCACTTGATGGGATCTTTCCAGAGCAGATGAGCCATTAACAGTAT 1062
QY 301 ProCysGln 303


```

/db_xref="taxon:9606"
/cclone="pBstAK13"
/tissue_type="lung"
/cclone_id="Lambda gct11"
163.1983

```

```

/product="TGF-beta activated kinase 1b"
/protein.id="BAA25026.1"
/db.xref="GI:2924626"
/translation="MSTSAASSSSSSSGAGMIAIPSYLFIETDIKEIIVEVVGGRGAGFVCAKMRKADVAIKQIESSEKRAFLVEIRQLSRVNHPIVAKLYGCALNPVCLVMEVAEGSSLVNLHGAPRLPYITPAAHMSICLQSOVAALHSMQAPALVHRDLKRPENLLVAAGTVLKIICFGTACDILQTHMTHMSAAMVMEVEGSSNSKRCDFVWGJILMVEYTRRRPFDEIGGPAFRIMAAVHNSTRPPLIKNLPRIESTLMTCSNDQPORPSEMEIVKIMTHLMRFPFGADEPLPFCVYDGSQNSAISTSPMDIATWMSNDSTNTPNMEOVATNDITIKRESKLKNOAKOQSSGSLISGASGSSVEELPSPSEKSSAMSEIARIIAATTAYTKRESKLKRRKAKKKAATFGNIIILPVEIYISGNQPPRRSIYDEKATGTPGVSSRSSPSVKNITTSQPTSEKTRSHPTPPDSTDITNSDMSIPMAVITLHQQLPLACPNKSEMAVFEQHKQAQETMKQVETIALLKQELVALLDQDEKQNTSRILVDEHKLLDNKSLSTYYQOQCKQVETIRSQOQKQOOTS"

```

Alignment Scores:	
Pred. No.:	4.64e-149
Score:	1615.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Matchn:	100.00%
DB:	9
	Gaps: 0
	Length: 2850
	Matches: 303
	Conservative: 0
	Mismatches: 0
	Indels: 0
	Gaps: 0

US-09-830-144-2_COPY_1_303 (1-303) X AB009357 (1-2850)

QY	1	MetSerThrAlaSerAlaAlaSerLeuSerSerSerSerSerAlaGlyMetLeu	20
Db	163	ATGTCTACAGCCTCGCCGCCCTCCCTCCCTCCTGCTTCGCGCGGATGATCCAA	222
QY	21	AlaProSerGlnValLeuAsnPhenGluGluLeaPylYrLysGluIleGluValGlu	40
Db	223	GCCCTTCCTCCAGTCTCAACTTTGAAAGATCGCTCAAGAGATCGAGGTGAAG	282
QY	41	ValValGlyArgGlyValaPheGlyValValCysLysAlaLysTrpArgAlaLysAspVal	60
Db	283	GTTGTGGAAAGAGAGCCTTTGGAACTTTTGGCAAGTAAAGTGAAGCAAAAGATTT	342
QY	61	AlaIleLysGlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln	80
Db	343	GCTATTAAACAATAGAAAGTGAATCTAGAGGAAGCCTTATTGTGAGCTTCGGCAG	402
QY	81	LeuSerArgValAsnHisProAsnIleValLysLeuYrGlyValaCysLeuAsnProVal	100
Db	403	TTATCCCGTGAACCATCTTAATGTATAAGCTTTATGGAGCCTGCTTGATCCAGTG	462
QY	101	CysLeuValMetGluYrAlaGluGlyGlySerLeuYrAsnValLeuHisGlyValaGlu	120
Db	463	TGCTTTGATGAAGATAAGTGAAGGGGCTTTATTAATATGCTCCAGTGGCTGAA	522
QY	121	ProLeuProYrYrThrAlaAlaHisAlaMetSerTrpCysLeuGlnCysSerGlnGly	140
Db	523	CCATTGCCATATTTACTGCTGCCCAACCCAAATGAGTTGTTTACAGTGTTCACAGA	582
QY	141	ValAlaYrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysProPro	160
Db	583	GTTGCTTATCTTCACAGATGCCAACCCAAACCGCTTAATTCACAGGAGACTGAACCA	642
QY	161	AsnLeuLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCys	180
Db	643	AACTTACTGCTGGTTGGCAGGGGACAGTTCTAAAAAATTTGTGATTTTGGTACAGCCTG	702
QY	181	AspIleGlnTrpHisMetThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGluVal	200
Db	703	GACATTAGACACATGATGACCAATTAACAAGGGAGTGTGCTTGATGATGGCACCTGAAGTT	762

QY 201 PheGluGlySerAsnIyrSerGluGlyCysAspValPheSerTrpGlyIleIleLeuTrp 220
Db 763 TTTGAAGGTAGTAATTACAGTGAATAATGTGACGCTTTCAGCTGGGGATTATTCCTTTGG 822
QY 221 GluValIleThrArgArgGlyProPheAspGluIleGlyGlyProAlaPheArgIleMet 240
Db 823 GAAGTGATTAACGGCTCGGAAACCTTTGATGATGATGTGGTGGCCAGCTTTCGAATCATG 882
QY 241 TrpAlaValHisAsnGlyThrArgProProLeuIleIleAsnLeuProGlyProIleGlu 260
Db 883 TGGGCTGTTCATATGTGTACTCGACCACTGATTAATAAAATTACTTAAGCCATTGAG 942
QY 261 SerLeuMetThrArgCysTrpSerIleAspProSerGluArgProSerMetGluGluIle 280
Db 943 AGCGTGATGATCGCTTTGTGTCTAAAGATCCCTTCCACGCGCCCTTCAATGAGAGAAATT 1002
QY 281 ValIleIleMetThrHisLeuMetArgTyrPheProGluIleAspGluProLeuGlnTyr 300
Db 1003 GTGAAAAATATGACTCATCTTGATGCGGTACTTTCAGAGACAGATGAGCATTACAGTAT 1062
QY 301 ProCysGln 303
Db 1063 CTTGTCTAG 1071

LOCUS	E38329	2866 bp	DNA	linear	PAT 31-JAN-2002
DEFINITION	NE-rappa B activation	inhibitory drug	targeting TAK1	and method for	identifying the same.
ACCESSION					

ORGANISM	unidentified
unclassified.	

TITLE
NF-kappa B activation inhibitory drug targeting TAK1 and method for identifying the same
Patent: JP 2000197500-A 4 18-JUL-2000;
JOURNAL CHIRUGYOKU

COMMENT	OS
	Unidentified
PN	JP 2000197500-A/4
PD	18-JUL-2000
PF	04-FEB-1999 JP 1999026603

```

PI MAOHISA SUGITA, HIROAKI SAKURAI, NORIKO KAGEYAMA, PI HIROSHI
HASEGAWA
PC C12Q1/48, A6IK31/00, A6IK31/00, A6IK45/00, C12N5/10, C12N9/99, PC
C12Q1/02,
PC G01N33/15, G01N33/50, G01N33/566//C12N15/09, C12Q1/68, (C12N15/09,
PC C12R1:91),
PC C12N5/00, C12N15/00, C12N15/00, C12R1:91)
CC Strandedness: Double;
CC Topology: Linear;
EH Key Location/Qualifiers
FT source 1. .2866
/organism='Unidentified'.

```

```

source      1. 2866
             /organism="unidentified"
             /db_xref="taxon:32644"
BASE COUNT  848 a 588 c 656 g 774 t
ORIGIN

```

Pred. No.:	4-67e-149	Length:	2666
Score:	1615.00	Matches:	303
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
GB:	6	Gaps:	0

QY 101 CysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnValLeuHisGlyAlaGlu 120
 DB 457 TGTCTTGATGATGAAATGACAGAGGGGGCTCATTGTATTAATGCTCAGTGCTGAA 516
 QY 121 ProLeuProTyrTyrThrAlaHisAlaMetSerTyrCysLeuGlnCysSerGlnGly 140
 DB 517 CCAATGCCCTTACTACACCTGCTCTCATGCGCAGACTGGTGTATTACAGTGTCCCAAGA 576
 QY 141 ValAlaTyrLeuHisSerMetGlnProLysAlaLeuLeuHisArgAspLeuLysProPro 160
 DB 577 GGGGCTTACTCAGACGATGACGCCCAAGCGCTGATTCACAGGAGACTCAAGCCTCCA 636
 QY 161 AsnLeuLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCys 180
 DB 637 AACTGCTGCTGCTGGTGCAGAGGAGCAGTTCTAAAATCTGGATTGGTATTCACGCTTGT 696
 QY 181 AspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyrMetAlaProGluVal 200
 DB 697 GACATCCAAACACACATGACCAATTAAGGAGAGTGGCTTGGATGGCGGCTGAAGTG 756
 QY 201 PheGluGlySerAsnTyrSerGlnLysCysAspValPheSerTyrGlyIleLeuLeuTyr 220
 DB 757 TTTGAAGGTAGACATATGAGTGAAGAGTGTGATGCTTCACTGGGAGTATTTATCCTCTGG 816
 QY 221 GluValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMet 240
 DB 817 GAAGTATTAACACGCCGGAACCTTCGATGAGATGGTGCCGACCTTCAGAAATCATG 876
 QY 241 TTPAlaValHisAsnGlyThrArgProProLeuIleLeuAsnLeuProLysProIleGlu 260
 DB 877 TGGGCTTCTCATATGACCTGCACCAACCATGATCAAAAATTTCACCTAAGCCCTTAGG 936
 QY 261 SerLeuMetThrArgCysTyrPheSerLysAspProSerGlnArgProSerMetGluGluIle 280
 DB 937 AGCTTATGATGACAGCTGTGGTGTAGTACCACTCATCTCAGCCCTTCATAGAGGAAAT 996
 QY 281 ValLysIleMetThrHisLeuMetArgLysPheProGlyAlaAspGluProLeuGlnTyr 300
 DB 997 GTGAAATATATGATCATCTGATGCTGCTCCGAGGAGGATGACCATTAACAGTAT 1056
 QY 301 ProCysGln 303
 DB 1057 CTTGTTCAG 1065
 RESULT 11
 LOCUS BC006665 3107 bp mRNA linear ROD 07-AUG-2002
 DEFINITION Mus musculus, clone MGC:5989 IMAGE:3499247, mRNA, complete cds.
 ACCESSION BC006665 GI:13879375
 VERSION BC006665.1 GI:13879375
 KEYWORDS house mouse.
 SOURCE Mus musculus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 3107)
 Strausberg, R.
 Direct Submission
 Submitted (27-APR-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Jeffrey Green M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu

Gunaratne, P.H., Garcia, A.M., Lu, X., HuiYk, S.W., Hale, S.M.,
 Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
 Richards, S., Gibbs, R.A.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
 Series: IRAX Plate: 9 Row: h Column: 9
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis, Genomescan gene prediction, Similarity but not identity
 to protein.
 FEATURES
 source
 1.3107
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /map="FVA/N"
 /clone="MGC:5989 IMAGE:3499247"
 /issue_type="Mammary tumor. C3(1)-Tag model. Infiltrating
 ductal carcinoma. 5 month old virgin mouse."
 /clone_lib="NCI CGAP_Mame"
 /lab_host="DH10B"
 /note="Vector: PCMV-SPORT6"
 148..1968
 /codon_start=1
 /product="Unknown (protein for MGC:5989)"
 /protein_id="AAH06665.1"
 /db_xref="GI:13879376"
 /translation="MSTASAASSSSSSSASEMIETAPSOVLNFEEDYIEVEEYVGR
 GAFGVCAKAKMKAQVVALKQISESEERAFIYELQLSRVNHPNIVKLYGACLNPLC
 VMEYAGSGSLNVNVLGAEPYPYTAAMMSWCLQCGOVALHSHQPYALHRLDKLP
 NILLVAGGTVLKICDFGPAODIOTHTNNKGSAAWAPVEFGSNVSEKDFVSGII
 LMEVITRRKPFDEIGPAPRIMWAVHNKTRPLIKLKPESILMTKMSKDPGSRPS
 MEIYKIMTHLMIRYPPGADEPIQYCCQVSDDEQNSATSGSFMDIATNTSKSDTN
 MEQVATNDITIRLSKLKNAQKQSEGRFSLDASRQSVESIPPTSEGRKMDM
 SEIEARIVATAVYSKPKRGHRTAGFNLDVPELVISGNGPARRSIQDLVTGTEP
 GGVSSRSSPSVPMITTSGETSEKPARSPHDDSTPDNGSDNSIPMAVYTLDDHQLQ
 PLAPCNSSKESMAVFEOHCKMAOEKVMYOTETIALILORKEQELVARELDQEKQQTSTR
 LVQEHKKLLDENKSLSTYQQCKKLEIVRSQQQKRGTS"
 BASE COUNT 898 a 687 c 772 g 750 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2e-148 Length: 3107
 Score: 1609.00 Matches: 302
 Percent Similarity: 99.67% Conservative: 0
 Best Local Similarity: 99.67% Mismatches: 1
 Query Match: 99.63% Indels: 0
 DB: 10 Gaps: 0
 US-09-830-144-2_COPY_1_303 (1-303) x BC006665 (1-3107)
 QY 1 MetSerThrAlaSerAlaAlaSerSerSerSerSerAlaGlyMetIleGlu 20
 DB 148 ATGTTCGACAGCCTCCGCCCTGCTCCTCCTGCTCTGCTCCAGTGAAGATCGAA 207
 QY 21 AlaProSerGlnValLeuAsnPheGluGluIleAspTyrLysGluIleGluValGlu 40
 DB 208 GCGCGCTGCAGGTCTGAACCTTGAAGAATCGACACAAAGAGATCGAGTGAAGAG 267
 QY 41 ValValGlyArgGlyAlaPheGlyValValCysLysValAlaTyrPArgAlaLysAspVal 60
 DB 268 GTTGTCCGAAGAAGAGACTTTTGGAGTAGTTTGCACCAACTAAGTGAAGACAAAAGATGC 327
 QY 61 AlaIleLysGlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
 DB 328 GCTATTAAACAGATAGAAAGTGAAGTCTGAGAGGAGGAGGCTTGTGGAGTCCCGCAG 387
 QY 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnProVal 100
 DB 388 TTGTGCGGTGTGAACCATCTCAACATTGTCAAGTTGACGAGGCTGCTGAATCCAGTA 447
 QY 101 CysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnValLeuHisGlyAlaGlu 120

```

Db 448 TGTCTTGTGATGGAATATGACAGAGGGGGCTCATTTGTTATATGTCGTGCTGCTGAA 507
Qy 121 ProLeuProTyrThrAlaHisAlaMetSerTyrCysLeuGlnCysSerGlnGly 140
Db 508 CATTCGCCCTTACTACTCTGCTCTATGCCATGAGCTGGTGTTCAGTGTTCCTCAAGA 567
Qy 141 ValAlaTyrLeuHisSerMetGlnProTyrAlaLeuLeuHisArgAspLeuLysPro 160
Db 568 GTGGCTTACTGTCACAGCATGACGCCAAGCGCTGATTTCACAGGACCTCAAGCTCCA 627
Qy 161 AnLeuLeuLeuValAlaGlyThrValLeuLysIleCysAspPheGlyThrAlaCys 180
Db 628 AACTTCTCTCTGTTGACAGAGGACAGTTCTAAATACTGCGATTTTGGTACACCTTGT 687
Qy 181 AspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyrMetAlaProGluVal 200
Db 688 GACATCCAAACACACATGACCAATAATAAAGGAGTGCTGCTTGGATGGCGCTGAAGTG 747
Qy 201 PheGluGlySerAsnTyrSerGlnLysCysAspValPheSerTyrGlyIleLeuTyr 220
Db 748 TTTGAAGGTAGCAATTACAGTGAAGAGTGTGATGCTTTCAGCTGGGGTATTATCTCTCG 807
Qy 221 GluValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMet 240
Db 808 GAAGTGATAACACGCCGGAACCTTCGATGAGATCGGTGGCCAGCTTTCAGAAATCATG 867
Qy 241 TrpAlaValHisAsnGlyThrArgProProLeuLeuLysAsnLeuProLysProIleGlu 260
Db 868 TGGGCTGTTCTAATGGCACTCGACCACCATGATCAAAATTTTACCTAAGCCCATTTGAG 927
Qy 261 SerLeuMetThrArgCysTyrSerLysAspProSerGlnArgProSerMetGluGluIle 280
Db 928 AGCTTGATGACAGCGTGTGGTCTAAGGACCATCTCAGCGCCCTTCAATGAGGAAT 987
Qy 281 ValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyr 300
Db 988 GTGAAATAATGACTCACTTGATGCGGTACTTCCAGGAGCGGATGAGCCATTACAGTAT 1047
Qy 301 ProCysGln 303
Db 1048 CCTTGTGAC 1056

RESULT 12
LOCUS XLU92030 2812 bp mRNA linear VRT 18-APR-1998
DEFINITION Xenopus laevis TGF-beta-activated kinase TAK1 mRNA, complete cds.
ACCESSION U92030
VERSION U92030.1 GI:3057035
KEYWORDS
SOURCE Xenopus laevis.
ORGANISM Xenopus laevis.
REFERENCE 1. (bases 1 to 2812)
AUTHORS Shibuya,H., Iwata,H., Masuyama,N., Gotoh,Y., Yamaguchi,K., Irie,K.,
Matsumoto,K., Nishida,E. and Ueno,N.
TITLE Role of TAK1 and TAB1 in BMP signaling in early Xenopus development
JOURNAL EMBO J. 17 (4), 1019-1028 (1998)
MEDLINE 98130593
PUBMED 9463380
REFERENCE 2. (bases 1 to 2812)
AUTHORS Shibuya,H.
TITLE Direct Submission
JOURNAL Submitted (05-MAR-1997) Faculty of Sciences, Hokkaido University,
Nishi 6-chome, Kita 12, Kita-ku, Sapporo, Hokkaido 060, Japan
FEATURES
source 1. 2812
/organism="Xenopus laevis"
/db_xref="taxon:8355"
341..2191
/notes="TGF-beta-activated kinase"
/codon_start=1
CDS

```

```

/product="TAK1"
/protein_id="AAC14008.1"
/db_xref="GI:3057036"
/tranlation="MSATSAEMIETPPVLNPEIDYKELEVEEVVGRGTFGVVCRAKW
RKQVAIKOIESESEKRAFIVELRQLSRVNHENIVKLYGACLNPLVCLWVEYAGGSLY
NLVHGAEPYIYTAHAHMSWLCQAGVAYLHSMKPKALIHRLDLPAPNLLVAGGTVL
KICDFTACDIQTHMTNNKSAAMWAEVFEESNYSEKCDVFSWGIILWEVITRRKPF
DEIGFAPFRIMWVHNGTRPPLIKNLKPKIESLMTRCWSKDPQPSMBEEIVKIMTHL
MOVFFGADVLOPCQYSDGOSGATSGCIDIITSTNSKSDIHIEPDPQASAS
NNTIKIESKLAHLKNOAKOTGESRSLSPSRGSSVESLSEIRGRPSTUGTSEGK
RMSADMSELEARISASTAVVVKPGRHKRTAFSGNILDVPEIITAGNQOQERSVDL
TIVTESOESNRSSRSPSVMITTTGTPDKPRGLWPAPDESSTNGSDNSIIPM
AYLTLDHQLQPLAPCNSKSWAEVFEQCKMAQYKMQVTEIALILQRKQELIAELDQ
DEKQONTSLRVQEHKKLLDENKSLSTYYQQCKQLEIVRSQQQRQGTs"
BASE COUNT 808 a 612 c 669 g 723 t
ORIGIN
Alignment Scores:
Pred. No.: 8,17e-137 Length: 2812
Score: 1490.50 Matches: 277
Percent Similarity: 96.92% Conservative: 6
Best Local Similarity: 94.86% Mismatches: 8
Query Match: 92.29% Indels: 1
DB: 5 Gaps: 1
US-09-830-144-2_COPY_1_303 (1-303) x XLU92030 (1-2812)
Qy 12 SerSerAlaGlyGluMetIleGluAlaProSerGlnValLeuAsnPheGluGluIle 31
Db 344 TCTCCACCTCTGCTGAAATGATTGAGACTCCA---CCGGTGTCTGAACCTTTGAGGAGATC 400
Qy 32 AspTyrLysGluIleGluValGluValGlyArgGlyAlaPheGlyValValCys 51
Db 401 GACTACAGGAGATTGAGGTGGAAGAGGTAGTTGGAAGAGAACCTTTGGTGTCTGTC 460
Qy 52 LysAlaLysTrpArgAlaLysAspValAlaIleLysGlnIleGluSerGluSerGluArg 71
Db 461 AAAGCTAAATGGCGAGGAAAGATGTGGCAATCAACAAATTTGAAAGCGAATCTGAAAGG 520
Qy 72 LysAlaPheIleValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLys 91
Db 521 AAAGCCCTTTATTGTGTGAGCTGCGACAGTGTCTGCTGTGTTAACCATCCCAATATTGTCAAG 580
Qy 92 LeuTyrGlyAlaCysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySer 111
Db 581 TTGTATGAGCGCTCCCTAAATCCGGTATGTTTAGTGTAGGATGTCAGAGGAGGTTC 640
Qy 112 LeuTyrAsnValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMet 131
Db 641 TTGTACAATGTTTTCATGGAGCTGAACCGCTGCCTTACTATACAGCTGCTCATGCAATG 700
Qy 132 SerTrpCysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAla 151
Db 701 AGTTGGTGTTTTACAATGTGCCAAGGAGTGTGCATATTTTACATAGCATGAAGCCGAGGCT 760
Qy 152 LeuIleHisArgAspLeuLysProAsnLeuLeuValAlaGlyGlyThrValLeu 171
Db 761 CTATTTACAGGAGACCTCAACACCACCAACTTGTCTGTGTAGCTGGAGGACCGTCTT 820
Qy 172 LysIleCysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGly 191
Db 821 AAGATTGTGACTTTGGTACAGCGCTGTGATATTTCAGACTCACACTGACTAATAACAAAGGA 880
Qy 192 SerAlaAlaTrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAsp 211
Db 881 AGTCAGCATGGATGGGTCCAGAAAGTTTTTGAAGGTAGCAATTACAGCGAGAAATGTGAC 940
Qy 212 ValPheSerTrpGlyIleIleLeuTyrGluValIleThrArgArgLysProPheAspGlu 231
Db 941 GTATTTAGTTGGGAAATAATTTCTTTGGAGAGTAAATAACCCGAGAAACCTTTCATGAA 1000
Qy 232 IleGlyGlyProAlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeu 251

```


Alignment Scores:		2.71e-127		Length: 135147	
Pred. No.:		1414.00		Matches: 270	
Score:		93.22%		Conservative: 5	
Best Local Similarity:		91.53%		Mismatches: 14	
Query Match:		87.55%		Indels: 7	
DB:		2		Gaps: 2	
US-09-830-144-2_COPY_1_303 (1-303) x AC114407 (1-135147)					
Qy	9	SerSerSerSerSerSerAlaGlyGluMetIleGluAlaProSerGlnValLeuAsnPhe	28		
Db	22102	TCTCCCTCTCTTGTCTTCTGGCAGTGAGATCAAAAGCGTGTCTGGGTCTGGAACATTC	22161		
Qy	29	GluGluIleAspTyrLysGluIleGluValGluValValGlyArgGlyAlaPheGly	48		
Db	22162	GAAGAGATTGACTACAGAGAGATCGAGGTGAAGAGGTTGTTGGAAGAGAGCTTTTGA	22221		
Qy	49	ValValCysLysAlaLysTrpArgAlaLysAspValAlaIleLysGlnIleGluSerGlu	68		
Db	22222	GTAGTTTGCAACGTAAGTGGAGCGAGCAAAAGATGTTGCTATTAAACAGATAGAAAGTCAG	22281		
Qy	69	SerGluArgLysAlaPheIleValGluLeuArgGlnLeuSerArgValAsnHisProAsn	88		
Db	22282	TCAGCAAGGAAGGCTTTCAATGTGGAGCTCCGACGTTGTCAGTGTGAACCATCTTAAT	22341		
Qy	89	IleValLysLeuTyrGlyValAlaCysLeuAsnProValCysLeuValMetGluTyrAlaGlu	108		
Db	22342	ATTGTAAGTTGATGAGCGCTCGCTGCATGCCAGTAGTCTTGATGAGTAATATACAGAG	22401		
Qy	109	GlyGlySerLeuTyrAsnValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaAla	128		
Db	22402	GGGGGCTCATTTGATAATGTGCTGCTGATGGTGTGTAACCATGCTTACTACACTGCTGCT	22461		
Qy	129	HisAlaMetSerTrpCysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGln	148		
Db	22462	CATCCAGTGAGCTGGTGTTTACAGCT-----TACCTGCACAGCATGCAG	22505		
Qy	149	ProLysAlaLeuIleHisArgAspLeuLysProProAsnLeuLeuValAlaGlyGly	168		
Db	22506	CCCAAGCTCTAATTCACAGGACCTCAAGCTTCCAAACTGCTGCTGGTGTGCAGGAGG	22565		
Qy	169	ThrValLeuLysIleCysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsn	188		
Db	22566	ACAGTTCTAAAAATCTCGGATTTTGATACAGCTTGTGACATCCAGACACACATGACCAAT	22625		
Qy	189	AsnLysGlySerAlaAlaTrpMetAlaProGluValPheGluGlySerAsnTyrSerGlu	208		
Db	22626	AATAAAGGAGTGTGCTGTGATGGTCCCCGAAGTGTGTTGAAGGTAGCAATATACAGTGAA	22685		
Qy	209	LysCysAspValPheSerTrpGlyIleIleLeuTrpGluValIleThrArgArgLysPro	228		
Db	22686	AAGTGTGATGCTTCACGTGGGGTATTATCTCTGGGAAGTGATAACAGCTGGAAACCC	22745		
Qy	229	PheAspGluIleGlyGlyProAlaPheArgIleMetTrpAlaValHisAsnGlyThrArg	248		
Db	22746	---GATGAGATCGTGGCCAGCTTTCAGATTATGAGACTGTTTCATTAATGGCACTGA	22802		
Qy	249	ProProLeuIleLysAsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSer	268		
Db	22803	CCACCACGTATCAAAAAATTTACCTAAGCCCATTTGAGAGCTTGATGACACACTGTTGGTCT	22862		
Qy	269	LysAspProSerGlnArgProSerMetGluGluIleValLysIleMetThrHisLeuMet	288		
Db	22863	AAGATCTCTTCTCAACGCCCTTCAATGGAGAAATTTGTGAAATAATATGACTCACTGTATG	22922		
Qy	289	ArgTyrPheProGlyAlaAspGluProLeuGlnTyrProCysGln	303		
Db	22923	CGGTACTTCCAGAGCGGATGAGCCATTACAGTATCTTGTCTGAC	22967		
RESULT 14					
AF199466					

LOCUS	AF199466	3349 bp	mRNA	linear	INV 17-APR-2000
DEFINITION	Drosophila melanogaster TGF-beta activated kinase 1 homolog mRNA, complete cds.				
ACCESSION	AF199466				
VERSION	AF199466				
KEYWORDS	GI:6289098				
SOURCE	Drosophila melanogaster.				
ORGANISM	Drosophila melanogaster. Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
REFERENCE	Takatsu, Y., Nakamura, M., Stapleton, M., Danos, M.C., Matsumoto, K., O'Connor, M.B., Shibuya, H. and Ueno, N. TAK1 participates in c-Jun N-terminal kinase signaling during Drosophila development Mol. Cell. Biol. 20 (9), 3015-3026 (2000)				
JOURNAL	Mol. Cell. Biol. 20 (9), 3015-3026 (2000)				
MEDLINE	20221548				
PUBMED	10757786				
REFERENCE	2 (bases 1 to 3349) Takatsu, Y., Nakamura, M., Stapleton, M., Danos, M., Matsumoto, M., O'Connor, M.B., Shibuya, H. and Ueno, N.				
AUTHORS	Direct Submission				
TITLE	Submitted (28-OCT-1999) Developmental Biology, National Institute for Basic Biology, 38 Nishigonaka, Myodaiji-cho, Okazaki, Aichi 444-8585, Japan				
JOURNAL	Developmental Biology, National Institute for Basic Biology, 38 Nishigonaka, Myodaiji-cho, Okazaki, Aichi 444-8585, Japan				
FEATURES	Location/Qualifiers				
source	1..3349 /organism="Drosophila melanogaster" /db_xref="taxon:7227" /chromosome="X" /map="19E5-7" 916..2952 /note="TAK1 homolog" /codon_start=1 /product="TGF-beta activated kinase 1 homolog" /protein_id="AAF06815.1" /db_xref="GI:6289099" /translation="MATASLDALAAAYVDFSEITLREKVGHSVVGVCKAVWRDKLVA VKEPFASAEOKDIEKVKQLSRVKHPNI IALHGISSVQQATYILIMEFAGSGSLHNF LHKVPAYSLAHAMSWARQCAEGLAYLHAMTPKPLIHRDVKPLNLLITNKRGNLIKDF GTVADKSTMTNNRSGSAAWAPEVFEKGYTEKCDIFSWAIVLWELSVLRKQPKFGIDN AYTKWIKYKGERPVLITCTPKRIEDLMTACWKTPEDEPSMQYIVGVHMEIVKDYTG ADKALETVFVNOQIVTAKESDCTVAAQPSLSOEGELSPSSQTLPTTAAANANVAIA ISKTTTSSMTENTSSTSDITPTNSGOLDNNPLPFVMTNRWDALPEESSESNSRDSFN LTSAAEATQRLETIRNGMLMACKPMEQLTLDVEANGFDELSPSESSSSSTNAKSDGRE RLTVDTKPYMMTIDLSNNNGSLHSHANGWQARDELQOEHEQELVNSLD VVDFDPEDKDTQESLAELIDPELQPEPPIPDNAESQLYLRDHRHMAKEYLSVHDNL YYAQDFDKDLIVQMDRTEREQKLELLRMKDKEGLQSLYNNLQQOQVASRQLAAGHPHQ PHPHPHNQLQHPHSHPMHFLQDEGGCLLPGSVCGGSESVGEVWVVPHPHNA"				
BASE COUNT	934 a 893 c 853 g 669 t				
ORIGIN					
Alignment Scores:					
Pred. No.:	1..54e-70		Length:		3349
Score:	818.00		Matches:		155
Percent Similarity:	69.93%		Conservative:		45
Best Local Similarity:	54.20%		Mismatches:		82
Query Match:	50.65%		Indels:		4
DB:	3		Gaps:		3
US-09-830-144-2_COPY_1_303 (1-303) x AF199466 (1-3349)					
Qy	17	GluMetIleGluAlaProSerGlnValLeuAsnPheGluGluIleAspTyrLysGluIle	36		
Db	913	GAGATGCCACAGCATCTCGTCGACGACATGCAGCGACGCTATGTGTGACTTCAGTGAGATA	972		
Qy	37	GluValGluGluValValGlyArgGlyAlaPheGlyValValCysLysAlaLysTrpArg	56		
Db	973	ACACTAAGAGAGAAAGTCTGGCCATGGGTCTTACGGTGTGGTCTGCAAGCGCGTTTGGCGC	1032		
Qy	57	AlaLysAspValAlaIleLysGlnIleGluSerGluSerGluArgLysAlaPheIleVal	76		
Db					

Db 1033 GACAAAGCTGTTCCGCTCAAGAGATTCTTCGCCAGCGCCGAGAGAGACATCGAAG 1092
Qy 77 GtluLeuArgGluLeuSerArgValAsnHisProAsnIleValIleuLeuTyrGly----- 94
Db 1093 GAGGTAAACAGATTGTTCGGCCGGTGAAGCACCCGAAACATATGCTCTGCACGGGATATCC 1152
Qy 95 AlaCysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsn 114
Db 1153 TCGTACCAGACAGCCACCTTACTGATTAATGAGATTGCGGAAAGGTGATGCTGCACAAAC 1212
Qy 115 ValLeuHisGlyValaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCys 134
Db 1213 TTCCTTCAAGGGC--AAGGTAAAGCCGGCATATTTCTGCGCCACGCGCATGAGCTGGGGC 1269
Qy 135 LeuGluCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHis 154
Db 1270 CGCAATGTCAGAGGGTCTGGCATTTTTCATGATGATGACCGCAAAACCATTAATACAT 1329
Qy 155 ATGAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCys 174
Db 1330 CGCGAGGTGAAGCCGCTGAACCTGCTCTTGACCAACAGAGAGCGCAATTCGAAGATATGC 1389
Qy 175 AspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAla 194
Db 1390 GACTTCGGCAGCGGTGGCGGGAAGTGCACATGATGATCCAAACATGCGCGTATGCGCGCT 1449
Qy 195 TrpMetAlaProGluValAlaPheGluGlySerAsnTyrSerGluLysCysAspValPheSer 214
Db 1450 TGGATGGCCGCCGAGGTCTTCGAAAGGCTCCAAGTATACGAGAAAGTGTACATTTTCAGC 1509
Qy 215 TrpGlyAlaIleLeuTyrGluValAlaIleThrArgArgLysProPheAspGluIleGlyGly 234
Db 1510 TGGGCCATTGTTTATAGGAGGTTCTGTCCAGAGACAGACCCCTTTAAAGGATTCGCAAT 1569
Qy 235 ProAlaPheArgIleMetTyrPalValHisAsnGlyThrArgProProLeuIleLysAsn 254
Db 1570 ---GCCTACACCATTCAGTGAAGGAACTCAAGAGGTGAACGCCCGCGCTGACCACT 1626
Qy 255 LeuProLysProIleGluSerLeuMetThrArgCysTyrSerLysAspProSerGlnArg 274
Db 1627 TGGCCCAAGCGCATCGAGACCTGATGACCGGCTGCGTGAAGAAACGGTGGCCAGAGATGCG 1686
Qy 275 ProSerMetGluGluIleValLysIleMetThrHisMetMetArgTyrPheProGluAla 294
Db 1687 CGCTCATGATGAGTACATGAGGCGCTTATGACAGAGATGTCAAGACTATACGGGGGCG 1746
Qy 295 AspGluProLeuGlnTyr 300
Db 1747 GACAAAGCCCTGGATATC 1764

RESULT 15
AY051953 3386 bp mRNA linear INV 27-AUG-2001
LOCUS Drosophila melanogaster LD42274 full length cDNA.
DEFINITION AY051953
ACCESSION AY051953.1 GI:15292216
VERSION FL1_CDNA.
KEYWORDS Drosophila melanogaster.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 3386)
Stapleton, M., Briskestein, P., Hong, L., Aghayani, A., Carlson, J.,
Champe, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George, R.,
Gonzalez, M., Guarin, H., Li, P., Liao, G., Miranda, A., Mungall, C. J.,
Munoo, J., Pacleb, J., Paragas, V., Park, S., Phouanavong, S., Wan, K.,
Yu, C., Lewis, S. E., Rubin, G. M. and Ceiniker, S.
Direct Submission
Submitted (10-AUG-2001) Berkeley Drosophila Genome Project,
Lawrence Berkeley National Laboratory, One Cyclotron Road,
Berkeley, CA 94720, USA
COMMENT
Sequence submitted by:

Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory
Berkeley, CA 94720
This clone was sequenced as part of a high-throughput process to
sequence clones from Drosophila Gene Collection 1 (Rubin et al.,
Science 2000) . The sequence has been subjected to integrity checks
for sequence accuracy, presence of a polyA tail and contiguity
within 100 kb in the genome. Thus we believe the sequence to
reflect accurately this particular cDNA clone. However, there are
artifacts associated with the generation of cDNA clones that may
have not been detected in our initial analyses such as internal
priming, priming from contaminating genomic DNA, retained introns
due to reverse transcription of unsplliced precursor RNAs, and
reverse transcriptase errors that result in single base changes.
For further information about this sequence, including its location
and relationship to other sequences, please visit our Web site
(<http://fruitfly.berkeley.edu>) or send email to
cdna@fruitfly.berkeley.edu.

FEATURES

source

Location/Qualifiers
1..3386
/organism="Drosophila melanogaster"
/strain="y; cn bw sp"
/db_xref="taxon:7227"
/map="19E1-19E1"
1..3386
/gene="CG18492"
/note="alignment with genomic scaffold AB003571"
/db_xref="FLYBASE:FBgn0042230"
935..2971
/gene="CG18492"
/note="Longest ORF"
/codon_start=1
/product="LD42274p"
/protein_id="AAK93377.1"
/db_xref="GI:15292217"

CDS

/translation="MATASLDALQAAVDFSSITLREKVGHGSYGVCVAKVARDKLVA
VKEFASAROKDLEKEVKQLSRVKNITIALGHSYQAOATYLIMEPFGSSLNKPIH
GKKRPAISLAAHMSAROCBGLATYIHMTPLRIHRDQKPLNLLITNKGRLKIDF
GVVADKSTMTNNRSGSAAMWAVEFEGSKYTECDI FSAVALMEVLSRKQPFKIDN
AVTIOKWKLYKGRSPRLITCPKRIEDLMTACWTVPEBRPSMQYIVGVNHELVKQYTG
ADKALRYTFVNOQIVTKESDGVAAQOPDLSQESGLSPSSQVITPTAANANVAIA
ISKTTSSMTENTSSTSDIPRNSGLODNNPLFLVMYRMRAALIEEESNBSRNSFN
LNSAARLRLRTIRNGKTIKNSNGGISHNSNLSHANGWQVRDELDQEHDEITVSLD
RLTVDTKRVMMTTDLSNNNGGISHNSNLSHANGWQVRDELDQEHDEITVSLD
VVDPDPDEDNDGTQSLAEIIDLPELQPEPIPDNAESQLIYRDRHMAKEYLSVDTNL
YVADQFKDLIVQMRTEREQKQELRLKKKDEKGLQSLYNNLQQQVARSOLAAGHHPO
PHRPHNPOLQHPHSHPMHFLQDRGCGLLPSGVCGSSVBEGMVVIPIPHNA"

BASE COUNT 956 a 898 c 857 g 675 t
ORIGIN

Alignment Scores:

Pred. No.: 1,56e-70 Length: 3386
Score: 818.00 Matches: 155
Percent Similarity: 69.93% Conservative: 45
Best Local Similarity: 54.20% Mismatches: 82
Query Match: 50.65% Indels: 4
DB: 3 Gaps: 3

US-09-830-144-2_COPY_1_303 (1-303) x AY051953 (1-3386)

Qy 17 GtluMetIleGluAlaProSerGlnValLeuAsnPhedGluIleAspTyrLysGluIle 36

Db 932 GAGATGGCCACAGATGCTGGAGCGCATCGACGACCTTATGACATTCGACGTGAGATA 991

Qy 37 GtluValGluGluValValGlyArgGlyAlaPheGlyValValCysLysValLysTyrParg 56

Db 992 ACATTAAGAGAAAGTCGCGCATGGGTCTTACGAGAGTGTCTGCAAGCCGTTTGGCGC 1051

Qy 57 AlaLysAspValAlaIleLysGlnIleGluSerGluSerGluArgLysAlaPheIleVal 76

Db 1052 GACAAAGCTGTTGCCGTCGCAAGAGATTCTTCGCCAGCGCCGAGAGAGACATCGAAG 1111


```
QY 77 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIysLeuTyrGly----- 94
Db 1112 GAGGTGAAGCAGTTGTCGCGCTGAAGCACCCGAACATCATCGCTCTGCACGGGATATCC 1171
QY 95 AlaCysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsn 114
Db 1172 TCGTACACAGCAGGCCACTTACTGATTAATGGAGTTCGCCGGAAGGTGGATCGTCGACAAAC 1231
QY 115 ValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCys 134
Db 1232 TTCCTTCACGGC--AAGGTGAAGCCGGCATATTCTCTGCCCCACGCCCATGAGCTGGCG 1288
QY 135 LeuGlnCysSerGlnGlyValAlaTyrIleHisSerMetGlnProLysAlaLeuIleHis 154
Db 1289 CGCCAATGTGCAGAGGGTCTGGCATATTTCATGCCATGACGCCAAAACCACTAATACAT 1348
QY 155 ArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCys 174
Db 1349 CGCAGCTGAAGCGCTGAACCTCTCTTGACCAACAGGACGCAATCTGAAGATATGC 1408
QY 175 AspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAla 194
Db 1409 GACTTCGGCAGCGTGGCGGACCAAGTCACCATGATGACCAACATCGCGCAGTGGCGT 1468
QY 195 TrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSer 214
Db 1469 TGGATGGCGCCGAGGTCTTCGAAGGCTCCAAGTATACGGAGAAGTGTGACATTTTAGC 1528
QY 215 TrpGlyIleIleLeuTyrGluValIleThrArgArgLysProPheAspGluIleGlyGly 234
Db 1529 TGGCCCATTTGTTATGGGAGGTTCCTCCAGGAAGCAGCCCTTAAAGGCATCGACAAT 1588
QY 235 ProAlaPheArgIleMetTyrAlaValHisAsnGlyThrArgProProLeuIleLysAsn 254
Db 1589 --GCCTACACCATCCAGTGGAGATCTACAAGGTGAACGCCCGCGCTGCTGACCACT 1645
QY 255 LeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArg 274
Db 1646 TGCCCCAAGCGCATCGAGGACCTGATACCGCCTGTGGAAAAACGGTGGCCGAGGATCGC 1705
QY 275 ProSerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAla 294
Db 1706 CCGTCGATGCAGTACATAGTGGCGTTATGCAGAGATCGTCAAGGACTATACGGGGCGC 1765
QY 295 AspGluProLeuGlnTyr 300
Db 1766 GACAAGGCCCTGGAATAC 1783
```

Search completed: December 10, 2002, 03:06:37

Job time : 3127 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 9, 2002, 23:14:54 ; Search time 326 Seconds
(without alignments)
2093.115 Million cell updates/sec

Title: US-09-830-144-2_COPY_1_303

Perfect score: 1615

Sequence: 1 MTTASASSSSSSSAGEMIE.....MTHLMRYFFGADPLQLYPCQ 303

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q/cgn2_1/USPTO.spool/US09830144/runat_04122002_141510_2901/app_query.fasta_1.455
-DB=N_Geneseq_101002 -QWTF=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=p2o -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USPR=US09830144 @CNG 1.1.79 runat_04122002_141510_2901 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_101002.*

1: /SID22/gcgdata/geneq/geneq-emb1/NA1980.DAT.*
2: /SID22/gcgdata/geneq/geneq-emb1/NA1981.DAT.*
3: /SID22/gcgdata/geneq/geneq-emb1/NA1982.DAT.*
4: /SID22/gcgdata/geneq/geneq-emb1/NA1983.DAT.*
5: /SID22/gcgdata/geneq/geneq-emb1/NA1984.DAT.*
6: /SID22/gcgdata/geneq/geneq-emb1/NA1985.DAT.*
7: /SID22/gcgdata/geneq/geneq-emb1/NA1986.DAT.*
8: /SID22/gcgdata/geneq/geneq-emb1/NA1987.DAT.*
9: /SID22/gcgdata/geneq/geneq-emb1/NA1988.DAT.*
10: /SID22/gcgdata/geneq/geneq-emb1/NA1989.DAT.*
11: /SID22/gcgdata/geneq/geneq-emb1/NA1990.DAT.*
12: /SID22/gcgdata/geneq/geneq-emb1/NA1991.DAT.*
13: /SID22/gcgdata/geneq/geneq-emb1/NA1992.DAT.*
14: /SID22/gcgdata/geneq/geneq-emb1/NA1993.DAT.*
15: /SID22/gcgdata/geneq/geneq-emb1/NA1994.DAT.*
16: /SID22/gcgdata/geneq/geneq-emb1/NA1995.DAT.*
17: /SID22/gcgdata/geneq/geneq-emb1/NA1996.DAT.*
18: /SID22/gcgdata/geneq/geneq-emb1/NA1997.DAT.*
19: /SID22/gcgdata/geneq/geneq-emb1/NA1998.DAT.*
20: /SID22/gcgdata/geneq/geneq-emb1/NA1999.DAT.*
21: /SID22/gcgdata/geneq/geneq-emb1/NA2000.DAT.*
22: /SID22/gcgdata/geneq/geneq-emb1/NA2001A.DAT.*
23: /SID22/gcgdata/geneq/geneq-emb1/NA2001B.DAT.*
24: /SID22/gcgdata/geneq/geneq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	1615	100.0	1704	20	AA99698	Human TGF-beta act	Human TGF-beta act
2	1615	100.0	1788	20	AA56285	Human TAK1-6xHis e	Human TAK1-6xHis e
3	1615	100.0	1959	18	AA785095	Human transforming	Human transforming
4	1615	100.0	2656	20	AA56279	Human TAK1 encodin	Human TAK1 encodin
5	1615	100.0	2656	21	AA39105	Human TAK-1 nucleo	Human TAK-1 nucleo
6	1615	100.0	2769	24	ABU88437	Pain regulated cDN	Pain regulated cDN
7	1615	100.0	2785	20	AA99696	Human TGF-beta act	Human TGF-beta act
8	1615	100.0	2866	20	AA99697	Human TGF-beta act	Human TGF-beta act
9	1609	99.6	2443	18	AA785094	Mouse transforming	Mouse transforming
10	818	50.7	3367	23	ABU02489	Drosophila melanog	Drosophila melanog
11	689.5	42.7	10997	23	ABU02488	Drosophila melanog	Drosophila melanog
12	505	31.3	759	23	ABL08337	Drosophila melanog	Drosophila melanog
13	482.5	29.9	3454	24	ABL70018	Pancreas cancer re	Pancreas cancer re
14	445	27.6	3141	22	AA18824	Human kinase (PKIN	Human kinase (PKIN
15	443	27.4	3538	24	AA34309	Human PKIN-12 cDNA	Human PKIN-12 cDNA
16	442	27.4	2120	21	AA299726	cDNA encoding huma	cDNA encoding huma
17	434.5	26.9	1224	21	AA43254	Arabidopsis thalia	Arabidopsis thalia
18	434	26.9	1063	22	AAH34976	Human colon cancer	Human colon cancer
19	434	26.9	1631	21	AA39537	Arabidopsis thalia	Arabidopsis thalia
20	434	26.9	1706	21	AAA75674	DNA encoding a hum	DNA encoding a hum
21	434	26.9	2191	22	AAH99263	Human protein enco	Human protein enco
22	434	26.9	2194	21	AA299734	Cardiovascular sys	Cardiovascular sys
23	434	26.9	2220	21	AA293783	Human survival reg	Human survival reg
24	434	26.9	2622	22	AAF75336	Human TGF-beta rec	Human TGF-beta rec
25	434	26.9	3967	22	AAH73366	Human cervical can	Human cervical can
26	433	26.8	1365	21	AA299727	cDNA encoding huma	cDNA encoding huma
27	433	26.8	2272	21	AA299736	Cardiovascular sys	Cardiovascular sys
28	433	26.8	2403	22	AA44701	Novel protein kina	Novel protein kina
29	432.5	26.8	3066	24	ABQ86185	cDNA encoding huma	cDNA encoding huma
30	432	26.7	2157	22	AAH46913	Novel human protei	Novel human protei
31	432	26.7	3111	24	ABN86357	Novel human protei	Novel human protei
32	432	26.7	3518	24	ABN86358	Arabidopsis thalia	Arabidopsis thalia
33	427.5	26.5	1591	21	AA40839	Arabidopsis thalia	Arabidopsis thalia
34	427.5	26.5	3072	23	ABU04355	Drosophila melanog	Drosophila melanog
35	419	25.9	3558	24	ABK83874	Human cDNA differe	Human cDNA differe
36	419	25.9	3558	24	AA361139	Human mitogen acti	Human mitogen acti
37	415.5	25.7	3063	23	ABU08336	Drosophila melanog	Drosophila melanog
38	411.5	25.5	2283	21	AA48526	Arabidopsis thalia	Arabidopsis thalia
39	410	25.4	2505	21	AAA47607	Human CARK (Cardia	Human CARK (Cardia
40	410	25.4	2505	21	AAA47609	Rat CARK (Cardiac	Rat CARK (Cardiac
41	410	25.4	2508	22	AAF44702	Novel protein kina	Novel protein kina
42	410	25.4	2788	23	AA580720	DNA encoding novel	DNA encoding novel
43	410	25.4	3025	21	AAA47606	Human CARK (Cardia	Human CARK (Cardia
44	410	25.4	3026	21	AAA47608	Rat CARK (Cardiac	Rat CARK (Cardiac
45	401	24.8	3365	24	ABK84203	Human cDNA differe	Human cDNA differe

ALIGNMENTS

RESULT 1

AA99698

ID AA99698 standard; cDNA to mRNA; 1704 BP.

XX AA99698;

XX 18-OCT-1999 (first entry)

DE Human TGF-beta activated kinase (TAK) 1c encoding nucleotide sequence.
XX Nuclear factor kappa B; NF-kB; inhibitor; TGF-beta activated kinase 1;
KW TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;
KW intractable disease; atrophic dermatitis; psoriasis; viral infection;
KW endotoxin shock; septicemia; human; hTAK1c; ss.

XX Homo sapiens.

XX Key

XX Location/Qualifiers

```

FT CDS 1..1704
FT /tag= a
FT /product= "hTAK1C"
PN WO9940202-A1.
PD 12-AUG-1999.
PF 02-FEB-1999; 99WO-JP00422.
XX
XX 30-OCT-1998; 98JP-0309316.
XX 06-FEB-1998; 98JP-0026003.
XX
XX (TANA ) TANABE SEIYAKU CO.
XX Hasegawa K, Kageyama N, Sakurai H, Sugita T;
XX
XX WPI; 1999-494298/41.
XX P-PSDB; AAY28998.
XX
XX Nuclear factor kappa B activation inhibitors, useful as preventives
XX for, e.g. autoimmune diseases
XX
XX Examples; Page 43-46; 49pp; Japanese.
XX
XX The invention provides a method for identifying or screening a nuclear
XX factor kappa B (NF-kB) activation inhibitor by examining the effect of a
XX test substance on modulating the function(s) of TGF-beta activated kinase
XX 1 (TAK1). The NFkB activation inhibitors targeting on TAK1 can be used to
XX treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis),
XX intractable diseases with inflammation (such as atrophic dermatitis and
XX psoriasis), viral infection, endotoxin shock, septicemia and others. The
XX present sequence represents the nucleotide sequence of human TAK1C
XX (hTAK1C) protein.
XX
SQ Sequence 1704 BP; 511 A; 381 C; 401 G; 411 T; 0 other;

Alignment Scores:
Pred. No.: 2.19e-173 Length: 1704
Score: 1615.00 Matches: 303
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 20 Gaps: 0

US-09-830-144-2_COPY_1_303 (1-303) x AAX99698 (1-1704)
QY 1 MetSerThrAlaSerAlaAlaSerSerSerSerSerAlaGlyGluMetIleGlu 20
DB 1 ATGTCTACAGCCTCTGCGCGCTCTCTCTCTCTCTGCGCGGTAGATGATCGAA 60
QY 21 AlaProSerGlnValLeuAnpHeGluGluIleAspTyrLysGluIleGluValGluGlu 40
DB 61 GCCCTTCCAGAGTCTCTCAACTTTGAAGAGATGACTACACAGAGATCAGAGTGAAGAG 120
QY 41 ValValGlyArgGlyAlaPheGlyValValCysLysAlaLysIlePargAlaLysAspVal 60
DB 121 GTTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
QY 61 AlaIleLysGluIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
DB 181 GCTATTAAACAATAGAAAGTGAATCTGAGAGAAAGCGTTTATTAGAGCTTCGCGAG 240
QY 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyValaCysLeuAnpProVal 100
DB 241 TTATCCCGTGTGAACCATCTATATTGTAAAGCTTTATAGAGAGCTGTGAATCCAGTG 300
QY 101 CysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGlyAlaGlu 120
DB 301 TGTCTGTGATGAAATATGCTGAAGGGGGCTTTTATATATATGCTGCATGCTGTA 360
QY 121 ProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeuGlnCysSerGlnGly 140

```

```

DB 361 CCATTGCCATATATATACTGCTGCCACGCAATGAGTGTGTTTACAGTGTCCCAAGA 420
QY 141 ValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysProPro 160
DB 421 GTGGCTTATCTTCACACAGACGCAACCCAAAGCGCTAATTACAGGAGCCGAAACACCA 480
QY 161 AsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCys 180
DB 481 AACTTACTGCTGTTGACAGGGGGGACAGTCTTAAAAATTTGTATTTTGTATACGCTGT 540
QY 181 AspIleGlnThrHisMetThrAsnAsnLysGlySerAlaIleThrMetAlaProGluVal 200
DB 541 GACATTCAGACACACATGACCAATTAACAAGGGAGTGCTCTTGATGATGCACTGAAGTT 600
QY 201 PheGluGlySerAsnTyrSerGluLysCysAspValPheSerTyrGlyIleIleLeuTyr 220
DB 601 TTTGAGGTAGTAAATTACAGTGAATAATGACGCTCTTCAGCTGGGGGTATATTCTTTGG 660
QY 221 GluValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMet 240
DB 661 GAAGTATTAACGCTCCGAAACCTTTGATGAGATTGGTGCCAGCTTCCGAATCATG 720
QY 241 TrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGlu 260
DB 721 TGGCTGTTCATATATGTTACTCCAGCACCACTGATAAATAATTACCTTAAGCCATTGAG 780
QY 261 SerLeuMetThrArgCysTyrSerLysAspProSerGlnArgProSerMetGluGluIle 280
DB 781 AGCTGTATGACTCTTGTGTGCTTAAAGATCTTCCAGCGCCCTTCATGAGAGGAAATT 840
QY 281 ValLysIleMetThrHisLeuMetArgTyrPheProGlyValaAspGluProLeuGlnTyr 300
DB 841 GTGAATAATATGACTCACTTATGCGGTACTTTCAGAGCAAGATGAGCAATTACAGTAT 900
QY 301 ProCysGln 303
DB 901 CTTGTCTAG 909

RESULT 2
AAX56285
ID AAX56285 standard; DNA; 1788 BP.
XX
XX AAX56285;
AC
XX
XX 21-JUL-1999 (first entry)
XX
XX Human TAK1-6xHis encoding DNA.
DE
XX
XX Human; TAB1; screening; inhibition; TGF-beta;
KW
XX transforming growth factor beta; ss.
XX
XX Homo sapiens.
OS
XX Synthetic.
XX
XX Key Location/Qualifiers
FT CDS /tag= a
FT
XX
XX WO9921010-A1.
XX
XX 29-APR-1999.
XX
XX 22-OCT-1998; 98WO-JP04796.
XX
XX 22-OCT-1997; 97JP-0290188.
XX
XX (CHUS ) CHUGAI SEIYAKU KK.
XX
XX Ohmoto T, Ono K, Tsuchiya M;
XX
XX WPI; 1999-312645/26.
XX
XX P-PSDB; AAY09547.
XX

```

PT Screening for TGF- beta inhibitory substances, which are useful as
PT drugs for treatment of diseases relating to its disorder

Example 1; Page 167-171; 195pp; Japanese.

A method has been developed for screening for substances which inhibit the binding of TAK1 polypeptide to TAK1 polypeptide. The method comprises: (a) contacting the polypeptide in the presence of a sample; and (b) detecting the amount of bound polypeptide, in which the sample can be pre-mixed with TAK1 or TAK1 polypeptide first. The transforming growth factor (TGF)-beta inhibitory substances can be used in drugs for indications e.g. as TGF-beta signal transduction inhibitors or activators, or extracellular matrix protein production enhancement inhibitors or activators, or cell proliferation prevention inhibitors or activators, or monocyte migration inhibitors or activators, or physiological activity induction inhibitors or activators, or immunosuppression inhibitors or activators, or amyloid beta protein precipitation inhibitors or activators, and such substances can also be inhibitors of the TAK1 polypeptide function, particularly kinase activity. The present sequence encodes TAK1-6xHis from an example of the present invention.

SQ Sequence 1788 BP; 551 A; 394 C; 413 G; 430 T; 0 other;

Alignment Scores:

Pred. No.:	2.35e-173	Length:	1788
Score:	1615.00	Matches:	303
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	20	Gaps:	0

US-09-830-144-2 COPY 1 303 (1-303) X AAX56285 (1-1788)

Qy	1	MetSerThrAlaSerAlaAlaSerSerSerSerSerSerAlaGlyGluMetIleGlu	20
Db	7	ATGTCACAGCCTGTGCGCGCTCTCTCTCTCTCTGCTTCGCGCGGTGAGATGATCGAA	66
Qy	21	AlaProSerGlnValLeuAsnPheGluIleAspTyrIlysGluIleValGluGlu	40
Db	67	GCCCTTCCACGGTCCTCAACTTTGAAGAGATCGACTCAAGGAGATCGAGTGTGAAGAG	126
Qy	41	ValValGlyArgGlyAlaPheGlyValValCysLysAlaLysTrpArgAlaLysAspVal	60
Db	127	GTGTGTGAAGAGAGAGCCTTTGGAGTGTGTTTGCAAAGCTAAGTGGAGAGCAAAAGATGT	186
Qy	61	AlaIleLysGlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln	80
Db	187	GCTATTAAACAAATAGAAAGTGAATCTGAGAGGAAAGCGTTTATTGTAGAGCTTCGGCAG	246
Qy	81	LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnProVal	100
Db	247	TTATCCCGTGTGAACCATCTTAATTTGTAAAGCTTTATGGAGCCTGCTTGAATCCAGTG	306
Qy	101	CysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGlyAlaGlu	120
Db	307	TGTCCTGTGATGAATATGCTGAAGGGGCTCTTTATATAATCTGCTGCATGTGTCTGAA	366
Qy	121	ProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeuGlnCysSerGlnGly	140
Db	367	CCATTTGCCATATTATATCTGCTGCCACGCATAGTTGGTGTGTTTACAGTGTTCCTCCAA	426
Qy	141	ValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgHisLeuLysProPro	160
Db	427	GTGCTTATCTTCACGATGCAACCCAAAGCGCTAATTCACAGGACCTGAAACCA	486
Qy	161	AsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCys	180
Db	487	AACCTTACTGTGTGTGAGGGGGACAGTCTCTAAAATTTGTGATTTGTGTACAGCCTGT	546
Qy	181	AspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGluVal	200
Db	547	GACATTCCAGACACATGACCAATAAACAGGGAGTCTCTCTTGGATGGACCTGAAGTT	606

Qy	201	PheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrpGlyllelleLeuTyr	220
Db	607	TTTGAAGGTAGTAATTACAGTGAATAATGTGACGTCTTCAGCTGGGGTATTATTCTTTGG	666
Qy	221	GluValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMet	240
Db	667	GAAGTGATAACGGGTGGAAACCTTTGATGAGATTGGTGGCCCGCTTTCCGAATCAIG	726
Qy	241	TrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGlu	260
Db	727	TGGGCTGTTCAATAATGGTACTCGACCACTCATATAAAAAATTTACCTAAGCCCATTTAG	786
Qy	261	SerLeuMetThrArgCysTrpSerLysAspProSerGlnArgProSerMetGluGluIle	280
Db	787	AGCCTGATGACTCGTTGTTGGTCTAAAGATCTCTTCCAGCGCCCTTCAATGGAGGAAAT	846
Qy	281	VallysIleMetThrHisLeuMetArgTyrPheProGlyValAlaAspGluProLeuGlnTyr	300
Db	847	GTGAAATAATGACTCACTTGATGCGGTACTTTCCAGGAGCAGATGAGCCATTACAGTAT	906
Qy	301	ProCysGln 303	
Db	907	CCTTGTCAG 915	
RESULT 3			
ID	AAT85095 standard; cDNA; 1959 bp.		
XX	AAT85095;		
XX			
DT	19-NOV-1997 (first entry)		
XX	Human transforming growth factor-beta activated kinase TAK-1 cDNA.		
XX	TGF-beta; signal transduction; TGF-beta activated kinase;		
KW	MAPK kinase activator; AMK-1; bone morphogenetic protein; BMP;		
KW	protein kinase; ss.		
XX	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	183..1922	
FT		/*tag= a	
FT		/product= TAK-1	
XX	JF09163990-A.		
XX	24-JUN-1997.		
XX	27-SEP-1996;	96JP-0256747.	
XX	24-JUL-1996;	96US-0685625.	
PR	29-SEP-1995;	95JP-0253549.	
XX	(CHUS) CHUGAI PHARM CO LTD.		
PA	(UENO/) UENO N.		
XX	WPI; 1997-380171/35.		
DR	P-PSDB; AAW27093.		
XX	DNA encoding transforming growth factor-beta-activated kinase, TAK-1		
PT	- useful for studying the TGF-beta signal transduction system		
XX	Claim 9; Page 13-15; 20pp; Japanese.		
XX	The present sequence encodes human transforming growth factor-beta		
CC	(TGF-beta) activated kinase, TAK-1. The DNA is used to produce the		
CC	TAK-1 protein which is involved in the TGF-beta family signal		
CC	transduction system. TAK-1, also known as activator of MAPK kinase		
CC	(AMK-1), is an enzyme which is activated by TGF-beta and bone		
CC	morphogenetic protein (BMP) and activates MAPK kinase by		

Db 363 GCTATTAAACAAATAGAAAGTGAATCTGAGAGAAAGCGTTATTGTGAGAGCTTCGGCAG 422
QY 81 LeuSerArgValAsnHisProAsnIleValIysLeuTyrglyAlaCysLeuAsnProVal 100
Db 423 TTAATCCCGTGTAACCATCTTAATTTGAAGCTTTATGAGCGCTGCTGATCAATG 482
QY 101 CysLeuValMetGluTyraIaGluGlySerLeuTyraAsnValIleHisGlyAlaGlu 120
Db 483 TGTCTTGATGGAATATCTGAAGGGGCGCTTTATATTAATGTCGTCATGTCCTGAA 542
QY 121 ProLeuProTyrrTyrrThraIalaHisAlaMetSerTrpCysLeuGlnCysSerGlnGly 140
Db 543 CCATTGCCATATTATATACGCTGCCACGCAATGAGTGGTGTTCACAGTGTCCCAAGGA 602
QY 141 ValAlaTyrrLeuHisSerMetGlnProIysAlaLeuIleHisArgAspLeuIysProPro 160
Db 603 GTGGCTTAATCTTCACAGCATGCAACCCAAAGCGCTTAATTCACAGGAGCCTGAAACCA 662
QY 161 AsnLeuLeuLeuValAlaGlyGlyThrValIleuLysIleCysAspPheGlyThrAlaCys 180
Db 663 AACTTACTGCTGGTTGCAGGGGGACAGTTCTAAAAATTGTGATTTTGTACAGCCTGT 722
QY 181 AspIleGlnThrHisSerMetThrAsnLeuIysGlySerAlaIleTrpMetAlaProGluVal 200
Db 723 GACATTCAACACACATGACCAATTAACAAGGGAGTGTGCTTGGATGGACCTGAAGTT 782
QY 201 PheGluGlySerAsnTyrrSerGluIysCysAspValPheSerTrpGlyIleIleLeuTrp 220
Db 783 TTTGAAGGTAGTATTAACGTAAAAATGTGACGTTCTGAGCGGGGTATTTCTTTGG 842
QY 221 GluValIleThrArgArgIysProPheAspGluIleGlyGlyProAlaPheArgIleMet 240
Db 843 GAAGTATTAACCGCTGGAAACCTTTGATGAGATGGTGGCCCAAGCTTTCCGAATCATG 902
QY 241 TrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeuProIysProIleGlu 260
Db 903 TGGCGTTTCATATGATGCTGACCAACCATGATAAAAATTTTACTTAAGCCCATTTGAG 962
QY 261 SerLeuMetThrArgCysTrpSerIysAspProSerGlnArgProSerMetGluGluIle 280
Db 963 AGCTGATGATGCTGTTGGTGTCTAAGATCCTTCCACGCGCCCTCAATGAGGAAATT 1022
QY 281 ValIysIleMetThrHisIleuMetArgIyrPheProGlyAlaAspGluProLeuGlnTyrr 300
Db 1023 GTGAAATATATGACTGACTTGATGCGGTACTTCCAGAGCAGATGAGCCATTACAGTAT 1082
QY 301 ProCysGln 303
Db 1083 CTTGTTCAG 1091

RESULT 6
ABL88437
ID ABL88437 standard; cDNA; 2769 BP.
XX ABL88437;
AC
XX
XX 16-MAY-2002 (first entry)
DT
XX
DE Pain regulated cDNA sequence 80.
XX
XX Pain; analgesic; gene therapy; neurological disorder;
KW neurodegenerative disease; gene; ss.
XX
XX Homo sapiens.
XX
XX MO200212338-A2.
PN
XX
XX 14-FEB-2002.
PD
XX
PF 03-AUG-2001; 2001WO-EP09011.
XX
XX 03-AUG-2000; 2000DB-1037759.
PR
XX

PA (CHEF) GRUENTHAL GMBH.
XX
XX Gillen C, Wezels I, Wendt S, Weine E, Schaefer MK;
PI
XX
XX WPI; 2002-257469/30.
DR
XX P-FSD; ABB85033.
PT
XX
XX Identifying pain-regulating compounds, useful for treating chronic pain
PT and for diagnosis, by measuring binding of compounds to specific
PT peptides and proteins
XX
XX
PS Claim 1; Fig 44; 213pp; German.
XX
XX The invention relates to identifying pain-regulating substances (A)
CC comprises (1) incubating a test substance with a cell (or preparation
CC from it) that has synthesised a peptide or protein (B) and (11) measuring
CC either binding of the test substance to (B) or some functional parameter
CC that is altered by this binding. The method is useful for identifying
CC pain-regulating substances (A) with analgesic activity. (A) along with
CC nucleic acid (ABL88411-ABL88441) that encode proteins (B).
CC ABB85006-ABB85037) that interact with (A); (B); vectors containing the
CC nucleic acid; antibodies against (B); cells that express (B) and agents
CC that bind to (B), are all useful for treating pain, particularly chronic
CC pain, including use in gene therapy. The same materials can also be used
CC for diagnosis, e.g. of neurological and neurodegenerative diseases. The
CC present sequence is that of a polynucleotide of the invention.
XX
SQ Sequence 2769 BP; 811 A; 565 C; 640 G; 753 T; 0 other;
XX
XX
XX Alignment Scores:
Pred. No.: 4,5e-173 Length: 2769
Score: 1635.00 Matches: 303
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
XX
US-09-830-144-2_COPY_1_303 (1-303) x ABL88437 (1-2769)
QY 1 MetSerThrAlaSerAlaAlaSerSerSerSerSerAlaGlyMetIleGlu 20
Db 163 ATGTCTACAGCTCTGCGCGCTCCTCTCTCGTTCGCGCGGTGAGATCGAA 222
QY 21 AlaProSerGlnValIleuAsnPheGluGluIleAspTyrrIysGluIleGluValIleGlu 40
Db 223 GCCCCTTCCAGGCTCCCACTTGAAGATGACCTACCAAGAGATCGAGTGGAAAG 282
QY 41 ValValGlyArgGlyAlaPheGlyValIaCysLysAlaIysTrpArgAlaIysAspVal 60
Db 283 GTTGTGGAAGAGAGCGCTTGGAGTGTGTTCCAAAGCTAAGTGAAGCAAAAGATGTT 342
QY 61 AlaIleYsgIleGluIleuAsnPheGluGluIleAspTyrrIysGluIleGluValIleGlu 80
Db 343 GCTATTAAACAAATAGAAAGTGAATCTGAGAGAAAGCGTTTATTGAGAGCTTCGGCAG 402
QY 81 LeuSerArgValaHisIleProAsnIleValIysLeuTyrrGlyAlaCysLeuAsnProVal 100
Db 403 TTAATCCCGTGGAACCATCTTAATATTGTAAGCTTTATGAGCGCTGGAATCCAGTG 462
QY 101 CysLeuValMetGluTyrrAlaGluGlySerLeuTyrrAsnValIleHisGlyAlaGlu 120
Db 463 TGTCTTGATGGAATATGCTGAAGGGGCGCTTTATATATATGCTGCAATGGTGGTGA 522
QY 121 ProLeuProTyrrTyrrThraIalaHisAlaMetSerTrpCysLeuGlnCysSerGlnGly 140
Db 523 CCATTGCCATATTATACGCTGCCACGCAATGAGTGGTGTTCACAGTGTCCCAAGGA 582
QY 141 ValAlaTyrrLeuHisSerMetGlnProIysAlaLeuIleHisArgAspLeuIysProPro 160
Db 583 GTGGCTTAATCTTCACAGATGCAACCCAAAGCGCTTAATTCACAGGAGCCTGAAACCA 642
QY 161 AsnLeuLeuValAlaGlyGlyThrValIleuLysIleCysAspPheGlyThrAlaCys 180

Qy	261	SeLseuMeTThArGcYstipserlyAspProserGlnArgProserMetGluGluLe	280
Db	943	ACCCTGATGACTCGCTTTGGTCTTAAACATCTTCCAGGCGCCCTTTAAATGAGGAAATT	1002
Qy	281	ValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyr	300
Db	1003	GTGAAATAATGACTCATCTTGATGGGACTTTCCAGAGAGATGAGCATTAAGATTT	1062
Qy	301	ProCysGln	303
Db	1063	CCCTGTCAG	1071
RESULT 8			
AAK99697			
ID	AAK99697	standard; cDNA to mRNA; 2866 BP.	
XX			
AC	AAK99697;		
XX			
DT	18-OCT-1999	(first entry)	
DE	Human TGF-beta activated kinase (TAK) 1b encoding nucleotide sequence.		
XX			
KW	Nuclear factor kappa B, NF-kB; inhibitor; TGF-beta activated kinase 1;		
KW	TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;		
KW	intractable disease; atrophic dermatitis; psoriasis; viral infection;		
KW	endotoxin shock; septicemia; human; hTAK1b; ss.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	163..1983	
FT		/*tag= a	
FT		/product= "hTAK1b"	
PN	W09940202-A1.		
XX			
PD	12-AUG-1999.		
XX			
PF	02-FEB-1999;	99WO-JP00422.	
XX			
PR	30-OCT-1998;	98JP-0309316.	
PR	06-FEB-1998;	98JP-0026003.	
XX			
PA	(TANA) TANABE SEIYAKU CO.		
XX			
PI	Haasegawa K, Kageyama N, Sakurai H, Sugita T;		
XX			
DR	WPI; 1999-494298/41.		
XX			
DR	P-PSDB; AAY28997.		
XX			
PT	Nuclear factor kappa B activation inhibitors, useful as preventives		
XX	for, e.g. autoimmune diseases		
XX			
PS	Examples; Page 39-43; 49pp; Japanese.		
XX			
CC	The invention provides a method for identifying or screening a nuclear		
CC	factor kappa B (NF-kB) activation inhibitor by examining the effect of a		
CC	test substance on modulating the function(s) of TGF-beta activated kinases		
CC	1 (TAK1). The NPKB activation inhibitors targeting on TAK1 can be used to		
CC	treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis),		
CC	intractable diseases with inflammation (such as atrophic dermatitis and		
CC	psoriasis), viral infection, endotoxin shock, septicemia and others. The		
CC	present sequence represents the nucleotide sequence of human TAK1b		
CC	(hTAK1b) protein.		
XX			
SO	Sequence 2866 BP; 848 A; 588 C; 656 G; 774 T; 0 other.		
SO	Alignment Scores:		

Alignment Scores:	
Pred. No.:	4,73e-13.
Score:	1615.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
	Length: 2866
	Matches: 303
	Conservative: 0
	Mismatches: 0

```

Query Match:      100.00%      Indels:      0
DB:               20          Gaps:         0
US-09-830-144-2_COPY_1_303 x AAAX99697 (1-2866)

```

[illegible]


```

XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656660/75.
XX P-PSDB; ABB58386.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
XX Claim 1; SEQ ID NO 1949; 21bp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
XX sequences (AB101840-AB116175) and the encoded proteins
XX (AB57737-AB572072).
XX
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 3367 BP; 938 A; 898 C; 856 G; 675 T; 0 other;
XX
Alignment Scores:
Pred. No.: 2.74e-82 Length: 3367
Score: 818.00 Matches: 155
Percent Similarity: 69.93% Conservative: 45
Best Local Similarity: 54.20% Mismatches: 82
Query Match: 50.65% Indels: 4
DB: Gaps: 3
US-09-830-144-2_COPY_1_303 (1-303) x ABL02489 (1-3367)
OY 17 GJumctileglualaproserglinvalleuasnphlegluglialeaspyrlysgluile 36
DB 932 GAGATGGCCACAGCATCGCTGGAGCCACTGCCAGCGACTTATGACTTACAGTACGATTA 991
OY 37 GIUValIGUValValValValValValValValValValValValValValValValVal 56
DB 992 ACACTAAGAGAGAAAGTGGCGCATGGCTCTACGAGAGTGGCTGGAGCCGTTGGCGC 1051
OY 57 AAlaYsAspValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 76
DB 1052 GACAAAGCTGTTGCCGTCAAGAGATTCTTCGCGACGCCGAGCAGAAAGACATCGAGAAG 1111
OY 77 GlueuArGlnleuSerArgValasnHisProasnleVallylsLeuTyrGly----- 94
DB 1112 GAGGTGAAGCAGATTGTCCGCGCTGAAGCACCAGAACATCATCGCTGCAGGATATCC 1171
OY 95 AAlaCysleuasnProValCysleuValMetGluTyrAlaGluGlyGlySerLeuTyrasn 114
DB 1172 TCGTACCGACAGGCCACCTACTGATATGAGATTCCGCGAAGGTGGATCCCTGCACAAAC 1231
OY 115 ValIeuHisGlyAlaGluProleuProTyrThrThrAlaAlaHisAlaMetSerTyrCys 134
DB 1232 TTCCTTCACGGC---AAGGTGAAGCCCGCATATTCCTCGGCCACGGCATGAGCTGGGCG 1288
OY 135 LeuGlnCysSerGlnGlyValAlaTyrIleuHisSerMetGlnProlySAlaLeuIleHis 154
DB 1289 CGCCAAATGTGACAGAGGCTGTGCATATTGTCATGCATGACGCCAATAACCATTAATCAT 1348
OY 155 ArgaspLeuIlyProProAsnleuLeuValAlaGlyGlyThrValLeuIlySileCys 174
DB 1349 CGCAGCGTGAAGCCGCTGAACCTCTCTTGAACCAAGAGCCCAATCTGAGAGATATGC 1408
OY 175 AspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnleuIlySGlySerAlaAla 194
DB 1409 GACTTCGGCAGCGTGGCGGACGACATGCATGATGACCAACATATCGCGGAGGCGGCT 1468
OY 195 TrpMetAlaProGluValPheGluGlySerAsnTyrSerGluIlyCysAspValPheSer 214

```

```

DB 1469 TGAATGGCGCCGAGGCTTCCTGAAAGCTCCAAAGTATACGAGAGAGTGTGACATTTTAC 1528
OY 215 TrpGlyIleIleLeuTyrGluValIleThrArgArgIlyProPheAspGluIleGlyGly 234
DB 1529 TGGGCACTTGTCTTAAGAGAGTTCTGTCCAGAGAGAGCCCTTTAAAGCATCGACAT 1588
OY 235 ProAlaPheArgIleMetTyrAlaValHisAsnGlyThrArgProProleuIleYsasn 254
DB 1589 ---GCTTACACATCCAGTGGAGATCTACAAAGGTGAAGCCCGCGCTGTCAGCACT 1645
OY 255 LeuProIlyProIleGluSerIleuMetThrArgCysTyrSerIlyAspProserGlnArg 274
DB 1646 TGCCCAAGGCGCATCGAGGACTGATGACCGCTGTGMAACGGTGGCCGAGATCCG 1705
OY 275 ProSerMetGluGluIleValIlySileMetThrHisIleuMetArgTyrPheProGlyAla 294
DB 1706 CCGTCGATGACATACATATGAGGCGCTTATGACGAGATGCTCAAGGACTATACGGGGCG 1765
OY 295 AspGluProLeuGlnTyr 300
DB 1766 GACAAAGCCCTGGAATAC 1783
RESULT 11
ID ABL02488 standard; cDNA; 10997 BP.
XX
XX ABL02488;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 1946.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ss.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
XX
XX 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656660/75.
XX P-PSDB; ABB58385.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
XX Claim 1; SEQ ID NO 1946; 21bp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
XX sequences (AB101840-AB116175) and the encoded proteins
XX (AB57737-AB572072).
XX
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 10997 BP; 3237 A; 2362 C; 2382 G; 3016 T; 0 other;
XX

```


CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilms' tumour.
XX
SQ Sequence 3454 BP; 594 A; 1217 C; 1136 G; 507 T; 0 other;

Alignment Scores:
Pred. No.: 4.14e-44 Length: 3454
Score: 482.50 Matches: 107
Percent Similarity: 56.63% Conservative: 51
Best Local Similarity: 38.35% Mismatches: 98
Query Match: 29.88% Indels: 23
DB: 24 Gaps: 7

US-09-830-144-2_COPY_1_303 (1-303) x ABL70018 (1-3454)

QY 21 AlaProSerGlnValLeuAsnPheGluGluIleAspTyrLysGluIleGluValGluGlu 40
Db 535 GCACCCGGGCGCTCCAGCTGCCCGAGAGATCCCTTCACGAGCTGCAGTAGAGGAG 594
QY 41 ValValGlyArgGlyAlaPheGlyValValCysLysAlaLysTyrArgAlaLysAspVal 60
Db 595 ATCATCGTGTGGGGCGCTTGGCAAGTCTATCGGCCCTGTGGCGTGGCGAGGAGTG 654
QY 61 AlaIleLysGlnIleGluSerGluSerGluArgLysAlaPheIleVal----- 76
Db 655 GCAGTCAAGCGCGCGCTGGAGCCTCGAAGAGGACCGCGCAGCTGCAGCGGAGCAGGTG 714
QY 77 -----GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGly 94
Db 715 TGCAGGAAGCGCGGCTCTTTGGAGCCCTGCAGCACCCCAACATAATTCGCCCTTAGGGGC 774
QY 95 AlaCysLeuAsnPro-----ValCysLeuValMetGluTyrAlaGluGlyGlySerLeu 112
Db 775 GCCTGCTCAACCCCGCCACACCTCTGCTAGTAGTGAGTATGCCCGGTGTGCACTG 834
QY 113 TyrAsnValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSer 132
Db 835 AGCAGGCTGTGGCAGGTGCGCGGTGCCACCTCACGTG-----CTGGTCAAC 882
QY 133 TrpCysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeu 152
Db 883 TGGGCTGTGAGGTGGCGCGGCGCATGAATACCTACACAATGATGCCCTGTGCCCATC 942
QY 153 IleHisArgAspLeuLysProProAsnLeuLeuVal----- 165
Db 943 ATCCACCGGACCTCAAGTCCATCACTCTGATCTCGAGGCCATCGAGACCAAC 1002
QY 166 AlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCysAsp-----IleGlnThr 184
Db 1003 CTCGACAGACCGGTGCTCAAGATCACGACTTCGGCTCGCCCGCAGTGGCACAGACC 1062
QY 185 HisMetThrAsnAnLysGlySerAlaAlaTrpMetAlaProGluValPheGluGlySer 204
Db 1063 ACCAAGATGAGCGCTGCGGGACCTACGCTGGATGCGCGGAGGTTATCCGCTCTCTCC 1122
QY 205 AsnTyrSerGluLysCysAspValPheSerTrpGlyIleIleLeuTrpGluValIleThr 224
Db 1123 CTCTTCTCCAAACAGTAGTGTGAGAGCTTCGGGTGCTGTGGAGCTGCTGAGC 1182
QY 225 ArgArgLysProPheAspGluIleGlyProAlaPheArgIleMetTrpAlaVal--- 243
Db 1183 GGGGAGGTCCCTACCGTGAGATC-----GACGCTTGGCGGTATGCGGTGGCT 1236
QY 244 HisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGluSerLeuMet 263
Db 1237 ATGAATAAGTACGCTGCCCATTCCTCCACGTGCGCGCGAGCCCTTTGCGCGCTCTCTG 1296

QY 264 ThrArgCysTrpSerLysAspProSerGlnArgProSerMetGluGluIleValLys 282
Db 1297 GAGGAATGCTGGGACCCAGACCCCGCGCGCGAGATTTCCGTTAGCATCTTGAAG 1353
RESULT 14
AAD18824
ID AAD18824 standard; cDNA; 3141 BP.
XX
AC AAD18824;
DT 18-DEC-2001 (first entry)
XX
DE Human kinase (PKIN)-9 cDNA.
XX
KW Human kinase; PKIN; gene therapy; adenocarcinoma; immune disorder; gout;
KW cancer; allergy; sarcoma; leukaemia; acquired immune deficiency syndrome;
KW AIDS; Addison's disease; microbial infection; inflammation; osteoporosis;
KW atherosclerosis; cardiovascular disease; myocardial infarction; anaemia;
KW myaesthesia gravis; cirrhosis; cataract; growth and development disorder;
KW seizure disorder; pulmonary embolism; Gaucher's disease; lipid disorder;
KW lipid storage disease; Pick's disease; Tay-Sachs disease; renal disease;
KW obesity; restorative therapy; immunomodulatory; vaccine; cardiovascular;
KW antimicrobial; cytostatic; antiinflammatory; asthma; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 1..3141
FT /*tag= a
FT /product= "Human PKIN-9 protein"
XX
PN WO200181555-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US12992.
XX
PR 20-APR-2000; 2000US-199021P.
PR 28-APR-2000; 2000US-200226P.
PR 05-MAY-2000; 2000US-202339P.
PR 11-MAY-2000; 2000US-203505P.
PR 18-MAY-2000; 2000US-205564P.
PR 26-MAY-2000; 2000US-207739P.
PR 01-JUN-2000; 2000US-208795P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Yue H, Gandhi AR, Tribouley CM, Kearney L, Griffin JA, Nguyen DB,
PI Bandman O, Lu DAM, Lal P, Burford N, Khan FA, Wallia NK, Yao MG;
PI Patterson C, Burrill JD, Marcus GA, Zingler KA, Recipon SA, Lu Y;
PI Policky JL, Thornton M, Tang YT, Hafalia A, Elliott VS, Baughn MR;
PI Walsh RT, Ramkumar J, Borowsky ML, Au-young J, Hillman JL;
PI Gururajan R;
XX
XX WPI; 2001-611740/70.
DR P-PSDB; AAE11775.
XX
XX Human kinases and nucleic acids, useful for preventing diagnosing and
PT treating cancers, inflammation and immune disorders -
XX
XX Claim 5; Page 157-158; 166pp; English.
XX
XX The present invention relates to human kinases (PKIN) and the nucleic
CC acids encoding them. PKIN is used as vaccine and in gene therapy. PKIN is
CC used in the prevention, diagnosis and treatment of diseases cancers,
CC adenocarcinoma, leukaemia, sarcoma, immune disorder, Addison's disease,
CC acquired immune deficiency syndrome (AIDS), anaemia, asthma, allergies,
CC gout, microbial infections, cardiovascular disease and/or inflammation,
CC myaesthesia gravis, atherosclerosis, cirrhosis, osteoporosis, myocardial
CC infarction, cataract, growth and development disorder, seizure disorder,
CC pulmonary embolism, Gaucher's disease, lipid disorder, lipid storage
CC disease, Pick's disease, Tay-Sachs disease, renal disease and obesity.
CC
CC PKIN may be used to treat disorders associated with decreased PKIN

expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of PKIN by expressing inactive proteins or to
 CC supplement the patient's own production of PKIN. PKIN nucleic acids may be
 CC used to produce the PKIN polypeptide, by inserting the nucleic acids into
 CC a host cell and culturing the cell to express the protein. PKIN nucleic
 CC acid and its complementary sequences may also be used as DNA probes in
 CC diagnostic assays to detect and quantitate the presence of similar
 CC nucleic acid sequences in samples and therefore which patients may be
 CC in need of restorative therapy. The present sequence is human PKIN-9
 CC cDNA.

XX
 XX Sequence 3141 BP; 685 A; 941 C; 942 G; 573 T; 0 other;

Alignment Scores:

Pred. No.:	Length:	3141
Score:	6,63e-40	111
Percent Similarity:	445.00	49
Best Local Similarity:	53.33%	107
Query Match:	27.55%	34
DB:	22	8
	Gaps:	

US-09-830-144-2_COPY_1_303 (1-303) x AAD18824 (1-3141)

Qy 3 ThrAlSerAlaAlaSerSerSerSerSerSerAlaGlyGluMetIleGluAlaPro 22
 Db 334 ACCCGCGGAGCGCCCTTCCAGCCCGCTGCCAGCCCGCGGCGAG----- 378
 Qy 23 SerGlnValIleuAsnPhgIuIleAspTyrLysGluIleGluValGluGluVal 42
 Db 379 -----GAGAAATTGATTGCGGAGCTCACCCTTGAAGAAGATTATT 420
 Qy 43 GlyArgGlyAlaPheGlyValValCysLysAlaLysTyrArgAlaLysValAlaIle 62
 Db 421 GGCATCGGGGGCTTGGGAAGGCTCTATCGTCTTCTGATAGGGGATGAGTGGCTG 480
 Qy 63 Lys-----GlnIleGluSerGluSerGluArgLysAlaPheIleValGluIleuArg 80
 Db 481 AAAGCACTGGCCAGCCCTGATGAGACATGACGCCAGACCAATAGAAATGTCGCCAA 540
 Qy 81 -----LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCys 96
 Db 541 GAGGCCAGCTTCCGCATGCTGGAAGCAACCCCAATCATTTGCCCTTAGAGGGGTATGT 600
 Qy 97 LeuAsnPro-----ValCysLeuValMetGlyTyrAlaGluGlyGlySerLeuTyrAsn 114
 Db 601 CTGAAGAGGCCCACTGCTGCTGATGAGATTGCTCTGAGAGACCTTTGAATAGA 660
 Qy 115 ValLeuHisGlyAlaGluProIleuProTyrTyrThrAlaAlaHisIleMetSerTyrCys 134
 Db 661 GTGTATCTGGGAAAAGATTCCCGCAGACATC-----CTGGTGAATTGGGCT 708
 Qy 135 LeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHis 154
 Db 709 GTCCAGATTGCCAGAGGAGTGAATCTTACTTGATGAGCAATTGTCCTCATCATCCAC 768
 Qy 155 ArgAspLeuLysProProAsnLeuLeuVal-----AlaGly 167
 Db 769 CGGACCTTAAGTCCAGCAACATATTGATCTCCAGAAAGTGGAGAATGAGACCTGAGC 828
 Qy 168 GlyThrValLeuLysTyrLeuAspPheGlyThrAlaCysAsp---IleGlnThrHisMet 186
 Db 829 AACCAAGATTCTGAAGATCTGATTTTGGCTGGCTGGGAATGGACCCAGCCCAAG 888
 Qy 187 ThrAsnAsnLysGlySerAlaAlaTyrMetAlaProGluValPheGluLysSerAsnTyr 206
 Db 889 ATAGATGCGCGGAGGACGATATGCTTGGATGGACCCGAAATCATCCGGGCTCCATGTTT 948
 Qy 207 SerGlnLysCysAspValPheSerTyrGlyIleIleLeuTyrGluValIleThrArgArg 226
 Db 949 TCCAAAGGAGCATGATGTGGAGACTATGGGGTGTCTACTTTGGAGATTGCTGACTGGTAG 1008
 Qy 227 LysProPheAspGluIleGlyGlyProAlaPheArgIleMetThrAlaValHisAsnGly 246

Db 1009 GTGCCCTTTCAGGCACTTGATGCTTACAGCT-CGC---TTATGAGATGGCCATGAAACAA 1064
 Qy 247 ThrArgPro-ProLeuIleLysAsnLeuProLysProIleGluSerLeuMetThrArgCys 266
 Db 1065 ACTGCGCCCTTCTATCTCTTCTTACAGTCCAGAACCTTTGGCAAACTATGGAAGCTG 1124
 Qy 266 sTrpSerLysAspProSerGlnArgProSerMetGluGluIleValLysIleMetThr 285
 Db 1125 CTGGAATCTGATCCCATCTCAGACACATCTTTCAGCAATATCTGTGACCACTAAC 1182

RESULT 15

ID AAD34309 standard; cDNA; 3538 BP.
 AC AAD34309;
 DT 16-JUL-2002 (first entry)
 XX
 DE Human PKIN-12 cDNA.
 XX
 KW Human; kinase; enzyme; PKIN-12 protein; immune system disorder; anaemia;
 KW acquired immune deficiency syndrome; thymic hypoplasia; Crohn's disease;
 KW asthma; neurological disorder; epilepsy; Charcot-Marie-Tooth disease;
 KW AIDS; seizures; cell proliferative disorder; cancer; adenocarcinoma;
 KW leukaemia; lymphoma; melanoma; myeloma; sarcoma; developmental disorder;
 KW Down's syndrome; gene therapy; protein therapy; cytoskeletal; gene; ss.
 XX Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..3294
 FT /*tag= a
 FT /product= "Human PKIN-12 protein"
 FT sig_peptide 1..51
 FT /*tag= b
 FT mat_peptide 52..3291
 FT /*tag= c
 FT /product= "Human mature PKIN-12 protein"
 XX
 PN W0200218557-A2.
 XX
 PD 07-MAR-2002.
 XX
 PF 31-AUG-2001; 2001WO-US27219.
 XX
 PR 31-AUG-2000; 2000US-229873P.
 PR 08-SEP-2000; 2000US-231357P.
 PR 14-SEP-2000; 2000US-232654P.
 PR 22-SEP-2000; 2000US-234902P.
 PR 29-SEP-2000; 2000US-236499P.
 PR 06-OCT-2000; 2000US-238369P.
 PR 13-OCT-2000; 2000US-240542P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Bandman O, Nguyen DB, Walla NK, Hafalia AUA, Yao MG, Gandhi AR,
 PI Gurunjan R, Ding L, Patterson C, Yue H, Baughn MR, Tribouley CM,
 PI Thornton M, Elliott JS, Lu Y, Jeon CH, Au-Young J, Tang YT,
 PI Azimzai Y, Burdill UD, Marcus GA, Zingler KA, Lu DM, Lal PG,
 PI Rankumar J, Warren BA, Kearney L, Policky JL, Thangavelu K,
 PI Burford N;
 XX
 DR WPI; 2002-329769/36.
 DR P-PSDB; AAE21717.
 XX
 PT New human kinases, useful for diagnosing, treating or preventing immune
 PT system disorders (e.g. Crohn's disease), neurological disorders (e.g.
 PT epilepsy), or cell proliferative disorders (e.g. cancers such as
 PT leukemia or lymphoma)
 XX
 PS Claim 91; Page 207-208; 218pp; English.
 XX
 CC The present invention relates to human kinases (PKIN) and polynucleotides

CC encoding such proteins. PKIN sequences of the invention are useful for
CC diagnosing, treating or preventing disorders associated with aberrant
CC expression of PKIN, particularly immune system disorders (e.g. acquired
CC immune deficiency syndrome (AIDS), thymic hypoplasia, Crohn's disease,
CC anaemia, asthma), neurological disorders (e.g. epilepsy, Charcot-Marie-
CC Tooth disease or seizures), cell proliferative disorders (e.g. cancers
CC such as adenocarcinoma, leukaemia, lymphoma, melanoma, sarcoma),
CC and developmental disorders (e.g. Down's syndrome). They are also used
CC in gene therapy and protein therapy. The present sequence is a cDNA
CC encoding human PKIN-12 protein.
XX

SQ Sequence 3538 BP; 763 A; 1055 C; 1062 G; 658 T; 0 other;

Alignment Scores:

Pred. No.:	1.34e-39	Length:	3538
Score:	443.00	Matches:	113
Percent Similarity:	53.97%	Conservative:	50
Best Local Similarity:	37.42%	Mismatches:	112
Query Match:	27.43%	Indels:	28
DB:	24	Gaps:	9

US-09-830-144-2_COPY_1_303 (1-303) x AAD34309 (1-3538)

QY	3	ThrAlaSerAlaAlaSerSerSerSerSerSerAlaGlyGlu-----MetIleGlu	20
DB	334	ACCCGCGCAGCGCCTTCTCCAGCGCTGCCAGCCGCGCGGAGGAGCCGCCAGTTGCTAC	393
QY	21	AlaProSerGlnValLeuAenPheGluGluIleAspTyrIysGluIleGluValGluGlu	40
DB	394	CGGCCCATTCAGTTGTTA-----GAAATTGATTTCGGGAGCTCACCTTGGGAAGAG	444
QY	41	ValValGlyArgGlyAlaPheGlyValValCysLysAlaLysTrpArgAlaLysAspVal	60
DB	445	ATTATTGGCATCGGGGCTTTGGGAAGGCTATCGTGCTTCGTGATAGGGATGAGGTT	504
QY	61	AlaIleLys-----GlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeu	78
DB	505	GCTGTGAAGCAGCTCGCCACGACCTGTATGAGGATCAGCCAGACCATAGAGAATGTT	564
QY	79	ArgGln-----LeuSerArgValAsnHisProAsnIleValLysLeuTyrGly	94
DB	565	CGCCAGAGCCCAAGCTCTTCGCCATGCTGAAGCAGCCCAACATCATTCCTTAAGAGGG	624
QY	95	AlaCysLeuAsnPro-----ValCysLeuValMetGluTyrAlaGluGlyGlySerLeu	112
DB	625	GTATGCTGAAGGAGCCCAACCTCTGCTTGTCATGGAGTTGCTCGTGGAGGACCTTTG	684
QY	113	TyrAsnValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSer	132
DB	685	AATAGAGTGTATCTGGGAAAAGGATTCCCCACAGATC-----CTGGTGAAT	732
QY	133	TrpCysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeu	152
DB	733	TGGGCTGTGCAGATTGCCAGGGGATGAATCTACTTGTATGAGCAATTTGTCCCATC	792
QY	153	IleHisArgAspLeuLysProProAsnLeuLeuVal-----	165
DB	793	ATCCACCGCGACCTTAAGTCCAGCACATATTGATCCTCCAGAGGTGGAGATGGAGAC	852
QY	166	AlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCysAsp---IleGlnThr	184
DB	853	CTGAGCAACAAGATCTGAAGATCACTGATTTTGGCCTGGGATGGCACCGCAACC	912
QY	185	HisMetThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGluValPheGluGlySer	204
DB	913	ACCAAGATGATCGGCGAGCGATGCTTGGATGGACCCCGAAGTCATCCGGGCTCC	972
QY	205	AsnTyrSerGluLysCysAspValPheSerTrpGlyIleIleLeuTrpGluValIleThr	224
DB	973	ATGTTTCCAAAGGCAGTGATGTGGAGCTATGGGTGCTACTTTGGGAGTTGCTGACT	1032
QY	225	ArgArgLysProPheAspGluIleGlyProAlaPheArgIleMetTrpAlaValHis	244

Db	1033	GGTGAGGTGCCCTTTCGAGGCAATTGATGGCTTAGCACT-CGC---TTATGGAGTGGCCAT	1088
QY	245	AsnGlyThrArgPro-ProLeuIleLysAsnLeuProLysProIleGluSerLeuMetTh	264
Db	1089	GAACAAACTCGCCCTTCTATTCCTTCTACGTGCCGAGAACCTTTTGCCTAACTCATGGA	1148
QY	264	rArgCysTrpSerLysAspProSerGlnArgProSerMetGluGluIleValLysIleMe	284
Db	1149	AGACTGCTGGAATCTGTATCCCACTCAGCACCATCTTTCACGAATATCTCTGGACCAGCT	1208
QY	284	tThr	285
Db	1209	AACC	1212

Search completed: December 10, 2002, 02:14:30

Job time : 341 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 10, 2002, 02:04:15 ; Search time 2206 Seconds
(without alignments)
2224.493 Million cell updates/sec

Title: US-09-830-144-2_COPY_1_303

Perfect score: 1615

Sequence: 1 MSTASAASSSSSSAGEMIE.....MTHLMRYFGADPELPYPCQ 303

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPIO_spo01/US09830144/runat_04122002_141511_2925/app.query.fasta_1.455
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09830144 @CGN 1 1 1716 @runat_04122002_141511_2925 -NCPU=6 -ICPU=3
-NO_XUPXY -NO_WMAP -LARGQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hcc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hcc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_nam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1379	85.4	1062	13	BM554120	BM554120 AGENCOURT
2	1242.5	76.9	998	9	AL525728	AL525728 AL525728
3	1167	72.3	739	13	BJ696710	BJ696710 603347214
4	1119	69.3	827	9	AL550589	AL550589 AL550589
5	1116	69.1	1006	14	BQ219348	BQ219348 AGENCOURT
6	1069	66.2	604	13	BM426610	BM426610 pgf2n.pk0
7	1063	65.8	929	9	AL520975	AL520975 AL520975
8	994.5	61.6	1054	14	BQ067693	BQ067693 AGENCOURT
9	984	60.9	646	13	BJ073883	BJ073883 BJ073883
10	982	60.8	688	13	BJ062988	BJ062988 BJ062988
11	982	60.8	696	13	BJ074867	BJ074867 BJ074867
12	953.5	59.0	910	12	BF780358	BF780358 602103276
13	921	57.0	747	13	BI334774	BI334774 602999677
14	920	57.0	542	10	AW960377	AW960377 EST372448
15	912	56.5	822	13	BI757369	BI757369 603025591
16	880.5	54.5	760	12	BG773488	BG773488 602720209
17	873	54.1	682	10	AV398933	AV398933 AV398933
18	852.5	52.8	745	13	BI767913	BI767913 603061057
19	843	52.2	681	10	AV398935	AV398935 AV398935
20	827.5	51.2	1191	13	EM472107	EM472107 AGENCOURT
21	823	51.0	641	14	BQ386875	BQ386875 NISC_mn20
22	814	50.4	1033	12	BE882182	BE882182 601505283
23	809	50.1	647	10	BB665645	BB665645 BB665645
24	794.5	49.2	594	13	BJ014293	BJ014293 BJ014293
25	748	46.3	686	12	BG548917	BG548917 602575750
26	741.5	45.9	585	9	AL773851	AL773851 AL773851
27	741	45.9	771	13	BI093821	BI093821 602860601
28	740.5	45.9	718	9	AL652289	AL652289 AL652289
29	720	44.6	516	13	BJ006787	BJ006787 BJ006787
30	715	44.3	586	13	BJ035129	BJ035129 BJ035129
31	714	44.2	605	9	AL704467	AL704467 DKF2p686A
32	712.5	44.1	634	9	AL710196	AL710196 DKF2p686A
33	712.5	44.1	700	10	BE291363	BE291363 601084896
34	704.5	43.6	831	14	BQ736206	BQ736206 AGENCOURT
35	689	42.7	845	12	BF027003	BF027003 601671216
36	681	42.2	642	10	BB645282	BB645282 BB645282
37	681	42.2	647	13	BJ008789	BJ008789 BJ008789
38	664.5	41.1	569	9	AUI77912	AUI77912 AUI77912
39	637	39.4	607	9	AI541706	AI541706 SD06739.5
40	609.5	37.7	644	9	AL782943	AL782943 AL782943
41	566.5	35.1	615	9	AL679013	AL679013 AL679013
42	557	34.5	858	13	BG916821	BG916821 602816079
43	553	34.2	711	10	AV718168	AV718168 AV718168
44	548	33.9	538	14	BQ557823	BQ557823 H4048E02-
45	532	32.9	447	9	AA306104	AA306104 EST177086

ALIGNMENTS

RESULT 1
BM554120
LOCUS
DEFINITION BM554120 1062 bp mRNA linear EST 20-FEB-2002
5', mRNA sequence.
ACCESSION BM554120
VERSION BM554120.1 GI:18793446
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1062)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

[illegible]

Db	122	CACAGTATGAAACC	AAGGCCCTTAATTCCACAGACACTGAACCAACAATTGGCTTG	181
OY	165	VAlAIAgIvGIyThrValleuysIIeCyAsapPhcglYThrAlaCysApIIEglInTr		184
Db	182	GTAAGCTGGGGGAC	CACTTTCAAAGACTGTGATTTGGTAGCGCTGTATATTCATAACA	241
OY	185	HISmetThrAsnAsnLysglYserAlaAlaIatPMeTalAProgluvalPhegluglYser		204
Db	242	CACATGACCAACAAT	AAGAGGAAGTGTGCTTGATGGACCGCATGAATTTTGAAGGTAC	301
OY	205	AnTyRSeEgLuILyScYasApVaIPheserTPrgLYlleleuITrgluValIIetHr		224
Db	302	AATTACAGAGAAAAA	AGTACGTTTTCACTTGGGGATTAATCTTTGGAGAGAAACAC	361
OY	225	ArGaAgLyPrOPheaspgluIIEglYgYP-roAlaPheArqIlmeITrpAlaVAlHis		244
Db	362	CCTAGGAACCTTTTGATATAGATGGTGTGCAGCTTTCGCATTAATGGGACGTTCAC		421
OY	245	AsngLYThArXProBLeuIlelySaenLeuProLySProlIEglUserLeuMetThr		264
Db	422	AATGGTACTCGACCA	CCACTGATCAAAAACCTTAACCAATGAGAGTTTAATGACC	481
OY	265	ArgCYstRseLysAspProSerGlnArgProSerMetGlunIleValylsIleMet		284
Db	482	CCCTGTGTGTC	CAAGGATCCCTCACACCACTTCATGAGAGAAATGTTAAATPATG	541
OY	285	ThriSleuMeTarGTYrPheProglYalAAspgluProleuGlnTyProcys		302
Db	542	ACACACTGATCGGTATCTTNNNAGCTGATGAACCTCTGACATCTTGC		595
RESULT 7				
AL520975				
LOCUS	ALS20975	LTI_NFL004_NB02	Homo sapiens cDNA clone CSDB002YM17	5
DEFINITION	prime, mRNA sequence.			
ACCESSION	AL520975			
VERSION	AL520975.1	GI:12784468		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1. (bases 1 to 929) Li W.B., Gruber C., Jeesee J. and Polayes D. Full-length cDNA libraries and normalization Unpublished (2001)			
AUTHORS	Contact : Genoscope			
JOURNAL	Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: segrel@genoscope.cns.fr, Web : www.genoscope.cns.fr.			
COMMENT	Location/Qualifiers			
FEATURES	Source			
	1..929	/organism="Homo sapiens"		
		/db_xref="taxon:9606"		
		/clone="CSDB002YM17"		
		/clone_id="LTI_NFL004_NB02"		
		/sex="male"		
		/tissue_type="neuroblastoma cells"		
		/lab_host="DH10B"		
		/note="Organ: brain; Vector: pCMVSORT 6; 1st strand cDNA was primed with a NotI-Oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetechn.com URL : http://fulllength.invitrogen.com"		
BASE COUNT	213 a	235 c	249 g	222 t
ORIGIN				10 others

Pred. No.:	4.33e-113	Length:	929
Score:	1063.00	Matches:	204
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.51%	Mismatches:	0
Query Match:	65.82%	Indels:	0
DB:	9	Gaps:	0

US-09-830-144-2_COPY_1_303 X AL520975 (1-929)

QY	1	MesSerThAlSeRlAlSaSeSseSseSseSseSseRlaIylImecIleglu	20
Db	305	ATGTCAAGCCTCGGCCGCTCCCTCCCTCCGTCGTTGGCCCGGTARATGATCGAA	364
QY	21	AlAProSeRglNValLeuAsnPhelugluNIleAspTYrlsGluilegluValglugu	40
Db	365	GCCCCCTTCCAGGTCTCACTACCTTGGAAGAATCACACAGAGATCCAGGTGGGAAGR	424
QY	41	VAlVaIGlVAyRGlVAlAPheGIvValVAlCySLysAlAlaYstTRprgrAlAlaYSapVal	60
Db	425	GTTGTTGGAAAGAGACCCTTGAGACTGTGTTGCACAACCTAGGGGACCAAAAAATGTT	484
QY	61	AlAlleYsgInlleGlusERglusERgluArglySApheIIeValgIuleuArgIn	80
Db	485	GCTATTAAACAATAGAAAGTAATCTGAGAGGAACCGTTTATTGTAGAGCTTGGGAG	544
QY	81	LeuSerArgValAsnHisProAsnIIeValLlysleuTYrGIyAlAcylsLeuAsnProVal	100
Db	545	TATCCCGGTGGAACCATCTTAATATTGTAAACCTTATAGAACCCGTGCTGAATCCAGG	604
QY	101	CySLeuValMeRglUTyAlAgIUglUGlYSerIEuTyASnValLeuHISglYlaglu	120
Db	605	TGCTCTTGATGGAAATATGCTGAAGGGGGCTCTTATATATATGCTGCATGCTGTAA	664
QY	121	ProLeuProTYrTYrThrAlAlAHISAlMeSeRTypCySLeugIncysSERingly	140
Db	665	CCATTGCCATATTATACTGCTGCCACAGCATAGATTGGTGTACAGTGTCCAAAGA	724
QY	141	VAlAlAYrLEuHSerMeTShProLYAlAlleuIHISArGSpeuLySProTo	160
Db	725	GTGGCTTATCTTCACAGCATCAACCCAAAAGCCCTAATTACAGGGACCTGAAACACCA	784
QY	161	AsnLeuLeuValAlagIyglYthrValLeuLysIEcysAsPheGIyThraIcys	180
Db	785	AACCTACGCTGGTGGAGGGGGACAGCTTAAAAATTGTGATTTGGTACACCTGT	844
QY	181	AspIIegInTHrHAmETThrasAsnLYSGlySErAlAlATrpMetAlaProGIuVal	200
Db	845	GACATTCAACAACAATACCAATMACMAAGGAGGTGCTGGATGCMCTGAATT	904
QY	201	PheGIuGLyseRAsn	205
Db	905	TTTGAAGGTAGTAAAT	919
RESULT 8	B0067693	1054 bp	mRNA linear EST 02-APR-2002
LOCUS	AGENCOURT 6643018	NIH_MGC_121	Homo sapiens cDNA clone IMAGE:5768401
DEFINITION	5' UENCL sequence.		
ACCESSION	B0067693		
VERSION	B0067693.1	GI:19896739	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
REFERENCE	1 (bases 1 to 1054)		
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabs-r@mail.nih.gov Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc.		

Oy	134	CysLeuGlcCysSerGlnGlyValAlaIleTyrLeuHISerMetGlnProLysAlaLeuIle	153
Db	63	TGTTTTCATCATGTGCCAAGGAGCTGCATTATTTTCATACGACTGAAGCCAAAGCGCTTGATT	122
Oy	154	HISArgAspLeuLysProPheAsnLeuLeuLeuValAlaGlyGlyThrValLeuLysIle	173
Db	123	CACAGGAGCACTCAAACCAACCACCAACTGGTGTGCTGTAAGCGAGGCACTGTTCTTAAGATT	182
Oy	174	CysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnLysGlySerAla	193
Db	183	TGTGACTTTGGTACACCGCTGTGATTTCAGACTCACATATACCTATTAACAAGAGTAGTCA	242
Oy	194	AlaTrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPhe	213
Db	243	GCAATGATGGCTCCACAGAAGTTTTTGAAGGTAGCAACTAACGCCAAAATAAGTGACGTGTT	302
Oy	214	SerTrpGlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGly	233
Db	303	AGTTGGGGCATTAATTTCTTTGGGAGTAATTAACCCGAAAGAAACCTTTCATGAAATTGCT	362
Oy	234	GlyProAlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLys	253
Db	363	GGTCCAGCGCTTCGGTAAATGTGGGCTGTTCACAAATGTACTCGGCCACCATTAATTAA	422
Oy	254	AsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGln	273
Db	423	AATTGGCTTAAGCTATTATTAAGCTTAATGACTCGCTGCTGGTCCMAAATCCCCACAA	482
Oy	274	ArgProSerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGly	293
Db	483	ACACCTTCAATGAGAGGAGATTGTCAAGATAAATGACACACATCTTAAGCAGATTCTTGGA	542
Oy	294	AlaAspGluProLeuGlnIntYrProCysGln	303
Db	543	GCAGACGTTTCCTTACAGTATCTCTTGTCAG	572
RESULT 10			
Bj062988			
LOCUS	Bj062988	688 bp mRNA linear EST 10-DEC-2001	
DEFINITION	Bj062988 NIBB Mochii normalized Xenopus tailbud library Xenopus laevis cDNA clone XL069ml7 5', mRNA sequence.		
ACCESSION	Bj062988		
VERSION	Bj062988.1 GI:17470746		
KEYWORDS	EST;		
SOURCE	Xenopus laevis		
ORGANISM	African clawed frog.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae; Xenopodinae; Xenopus.		
AUTHORS	Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-I,T. and Kohara,Y.		
TITLE	Expressed genes in X. laevis embryo		
JOURNAL	Unpublished (2001)		
COMMENT	Contact: Tadasu Shin-i Center For Genetic Resource Information National Institute Of Genetics 111 Yata, Mishima, Shizuoka 411-8540, Japan Tel: 81-559-81-6856 Fax: 81-559-81-6855 Email: tshin@genes.nig.ac.jp. Location/Qualifiers		
FEATURES			
source	1..688		

BASE COUNT	206 a	139 c	154 g	188 t	1 others
ORIGIN					
Alignment Scores:					
Pred. No.:	7.45e-104				668
Score:	982.00				177
Percent Similarity:	97.83%				Conservative: 3
Best Local Similarity:	96.20%				Mismatches: 4
Query Match:	60.80%				Indels: 0
DB:	13				Gaps: 0
US-09-830-144-2_COPY_1_303 (1-303) x BU062988 (1-688)					
QY	120	GLProleuProcyTrpYrThrAlaAlaHisAlaMetSerTrpCysLeuGlnCysSerGln	139		
Db	22	GAACCTTTGGCTTACTACTACTGCTGCCCATGCATAGATTGGATTACCAATGGCCCA	81		
QY	140	GLyAlaAlaTyrlleuHisSerMetGlnProlySAlaLeuIleHisArgAspLeuLyPro	159		
Db	82	GGAGTTGCATATTATTCATGACATGAAGCCAAAGGCTCTGATTCACAGGGACCTCAACCA	141		
QY	160	ProAsnLeuLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAla	179		
Db	142	CCAAACTTGTGTGGTGTACTGAGGACACTGTTCTTMAGATTGTGACCTTGTGTACAGCC	201		
QY	180	CysAspIleGlnThrHisMetThrAsnLeuLysGlySerAlaAlaTrpMetAlaProGlu	199		
Db	202	TGTGATATTACAGCTCAGCTACATGACTAATAACAAAGAAAGCAGCATGATGGCTCCAGAA	261		
QY	200	ValPheGluGlySerAsnTySerGlnLysCysAspValPheSerTrpGlyIleIleLeu	219		
Db	262	GTTTTTGAAGGTAGCAACACTACAGCCGAAAATGTGACGTTTAGTTGGGGCATTTATCTT	321		
QY	220	TrpGluValAlleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIle	239		
Db	322	TGGGAAGTAAATACCCGAGAAACCTTTGATGAAATGGTGGTCCACGTTCCGTATA	381		
QY	240	MetTrpAlaValAlaHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIle	259		
Db	382	ANGTGGGCTGTTCACAAATGTACTCGGCCACCATTTATTTAAATTTGGCTTAAGCTATT	441		
QY	260	GluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgProSerMetGluGlu	279		
Db	442	GAAAGCTTAATGACTCGCTGCTGTCCAAAGATCCCCCAAAAGACCTTCAATGGAGGAG	501		
QY	280	IleValLysIleMetThrHisLeuMetArgTrpPheProGlyAlaAspGluProLeuGln	299		
Db	502	ATTGTCAGATATATGACACATCTTAAGCAGATATTTTCTCGAGCAGACGTTCTTACAG	561		
QY	300	TyrProCysGln	303		
Db	562	TATCCTGTTCAG	573		
RESULT 11					
LOCUS	BU074867		696 bp	mRNA	linear EST 11-DEC-2001
DEFINITION	BU074867 NIBB Mochii normalized Xenopus calbind library Xenopus laevis cDNA clone XL071110 5', mRNA sequence.				
ACCESSION	BU074867				
VERSION	BU074867.1		GI:17505056		
KEYWORDS	EST.				
SOURCE	African clawed frog.				
ORGANISM	Xenopus laevis				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae; Xenopodinae; Xenopus.				
AUTHORS	Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara,Y.				
TITLE	Expressed genes in X. laevis embryo				
JOURNAL	Unpublished (2001)				

COMMENT

Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers

FEATURES

source
1..696
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="XL071110"
/clone_lib="NIBB Mochii normalized Xenopus tailbud library"
/tissue_type="whole embryo"
/dev_stage="stage 25"
/note="Vector: pBSRN3; Site 1: NotI; Site 2: EcoRI; cDNAs were oligo-dT primed and directionally cloned. Staging according to Nieuwkoop and Faber. Library is subtracted and was constructed by N. Garrett and A.M. Zorn, (Wellcome/CRC Institute)."

BASE COUNT 210 a 142 c 155 g 188 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 7.59e-104 Length: 696
Score: 982.00 Matches: 177
Percent Similarity: 97.83% Conservative: 3
Best Local Similarity: 96.20% Mismatches: 4
Query Match: 60.80% Indels: 0
DB: 13 Gaps: 0

US-09-830-144-2_COPY_1_303 (1-303) x BJ074867 (1-696)

QY 120 GluProLeuProTyrThrAlaAlaHisAlaMetSerTrpCysLeuGlnCysSerGln 139
DB 22 GAACCTTTGGCTTACTATACCTGCTGCCATGCAATGAGTTGGTTTACAAATGTCGCCAA 81
QY 140 GlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysPro 159
DB 82 GGAGTTGCATATTACATACATGAGCCAAAGCTCTGATTCACAGGACCTCAAAACCA 141
QY 160 ProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAla 179
DB 142 CCAAACTTTGTTGGTAGCTGGAGGCACTGTTCTTAAGATTGTCACTTTGGTACAGCC 201
QY 180 CysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGlu 199
DB 202 TGTGATATTACAGACTCACATGACTAATAACAAAGGAAGTGCAGCATGGATGGCTCCAGAA 261
QY 200 ValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrpGlyIleIleLeu 219
DB 262 GTTTTGAAGGTAGCAACTACAGCAAAATGTGACGTGTTAGTTGGGGCATATTCTT 321
QY 220 TrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIle 239
DB 322 TGGGAAGTAATAACCCGAAGAAACCTTTTCGATGAAATTTGGTCCAGCGTTCCGTATA 381
QY 240 MetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIle 259
DB 382 ATGFGGCGTGTTCACATGGTACTCGGCCACCATTAATTAATAAATTTGGCTAAGCCATT 441
QY 260 GluSerLeuMetThrArgCysTTrpSerLysAspProSerGlnArgProSerMetGluGlu 279
DB 442 GAAAGCTTAATGACTCGCTGCTGGTCCAAAGATCCCCCAAGACCTTCAATGGAGGAG 501
QY 280 IleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGln 299
DB 502 ATGTCAAGATAATGACATCATCAAGCAGTATTTTCTCGAGCAGACGTTTCTCTTACAG 561
QY 300 TyrProCysGln 303
DB 562 TATCCTTGTTCAG 573

RESULT 12

BF780358
LOCUS 602103276F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:4221379
DEFINITION 5', mRNA sequence.
ACCESSION BF780358
VERSION BF780358.1 GI:12085481
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 910)
NIH-MGC http://mgc.nci.nih.gov/.
NATIONAL INSTITUTES OF HEALTH, Mammalian Gene Collection (MGC)
UNPUBLISHED (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLAM9806 row: h column: 20
High quality sequence stop: 664.

FEATURES

source
1..910
Location/Qualifiers
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4221379"
/clone_lib="NCI_CGAP_Kid14"
/lab_host="DH10B (TI phage-resistant)"
/notes="Organ: kidney; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Constructed by Life Technologies. Note: this is a NCI CGAP Library. |"
BASE COUNT 247 a 209 c 246 g 207 t 1 others
ORIGIN

Alignment Scores:

Pred. No.: 2.45e-100 Length: 910
Score: 953.50 Matches: 194
Percent Similarity: 92.86% Conservative: 1
Best Local Similarity: 92.38% Mismatches: 6
Query Match: 59.04% Indels: 12
DB: 12 Gaps: 1

US-09-830-144-2_COPY_1_303 (1-303) x BF780358 (1-910)

QY 103 ValMetGluTyrAlaGluGlyGlySer-Leu-----TyrAsnVa 115
DB 4 GTGATGGAATATGCAGAGGGGGCTCATGTAATATGTTGTCGTCCTTCTCTACAGT 63
QY 115 lIeUHiSGlyValAlaGluProLeuProTyrThrAlaAlaHisAlaMetSerTrpCysLe 135
DB 64 GCTGCATGGTGTGAACCATTTGCCCTTACTACACTGCTCTCATGCCATGAGCTGGTGT 123
QY 135 uGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisAr 155
DB 124 ACAGTGTTCACAGAGTGGCTTACCTGCACAGCATGCAGCCCAAGCGCTGATTCACAG 183
QY 155 gAspLeuLysProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAs 175
DB 184 GGACCTCAAGCTCCAAACTTGTGCTGTTGCAGAGGGGACAGTCTTAAATAATCTGCGA 243
QY 175 pPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTr 195
DB 244 TTTTGGTACAGTTGTGATCCTCAACACACATGACCAATAATAAGGGAGGTGCTGCTTG 303

QY	195	pncet1apProGluVal1PheGluG15er1snYrSerG1u1y5CY5ASVal1PheSer1r	215
Db	304	GATGGCGCGCTGAAGTGTGGAAGGATACGATTAACAGTGAAGAAAAGTGTGATGCTTCACGCG	363
QY	215	PG1y11e11e1e1u1rPG1uVal11e-1Th1ArGArG1y5Pr0PheASg1u11eG1yG1yP	235
Db	364	GGGTATTACCTCCTGGGAAGTATTAACAACGCCGGAAACCTTCGATGAGATCGGTGGCC	423
QY	235	roA1PheArG11eMeC1rPa1aVal1H1ASnG1Y1Th1r1rPr0P1eAu11e1yASn1	255
Db	424	CAGCTTTCGAATCATGTGGCTGTTCATTAATGGCACTGACCACTCATGATCAAAAATT	483
QY	255	eup1r0Ly5Pr011eG1u5er1e1uMe1rTh1ArG1y5Tr1pSer1y5aPPr0SerG1nArGp	275
Db	484	TACCTAAGCCCTTGAGACCTGATGACACGCTG-TGGCTTAAGAGCCATCTACGCG-C	541
QY	275	ro5erMe1G1u1u11eVal1y511eMe1rTh1H151e1uMe1rG1y1rPr0G1yAlA	295
Db	542	CTTCAATGAGGAATGTAAGAAATATATACCTGATGAGGATCTTCCAGAGAGCGG	601
QY	295	SPG1uPr0LeuG1n1YrPr0CY5G1n	303
Db	602	ATGAGCA-TTACAGTATACTTCTCAG	626
RESULT 13			
LOCUS	B1334774	747 bp	mRNA linear EST 30-JUL-2001
DEFINITION	6029996777P1 NIH_MGC_12 Homo sapiens CDNA clone IMAGE:5141851 5',		
ACCESSION	B1334774		
VERSION	B1334774.1	GI:15019431	
SOURCE	EST.		
KEYWORDS	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Euhetia; Primates; Carnivora; Homnidae; Homo.		
TITLE	1 (bases 1 to 747)		
JOURNAL	NIH-MGC http://mgs.nci.nih.gov/ .		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgabbs-remail.nih.gov		
	Tissue Procurement: ATCC		
	CDNA Library Preparation: Life Technologies, Inc.		
	CDNA Library Arrayed by: Incyte Genomics, Inc.		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov		
	Plate: LAM11349 row: 1 column: 20		
	High quality sequence stop: 740.		
FEATURES	Location/Qualifiers		
source	1..747		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5141851"		
	/clone_lib="NIH MGC 12"		
	/r1sue_type="cervical carcinoma cell line"		
	/lab_host="DH10B"		
	/note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.4 kb. Library prepared by Life Technologies"		
BASE COUNT	173 a 187 c 214 g 173 t		
ORIGIN			
Alignment Scores:			
Pred. NO.:	1.1e-96	Length:	747
Score:	921.00	Matches:	186
Percent Similarity:	97.41%	Conservative:	2
Best Local Similarity:	96.37%	Mismatch:	3
Query Match:	57.03%	Indels:	0
DB:	13	Gaps:	3

UN-09-830-144-2_COPY_1_303 (1-303) x B13344774 (1-747)
 QY 1 MetSerThrAlaSerAlaAlaSerSerSerSerSerSerAlaGlyLmetLleGlu 20
 Db 174 ATGCTTACAGCCTTCGCCGCTCCCTCCCTCCTGCTTGGCGGAGATGATCCAA 233
 QY 21 AlaProSerGlnValIleuAsnPheGluGluIleAspTyrLysGluIleGluValGluGlu 40
 Db 234 GCCCTTCCAGGCTCCTCACTTTGAAAGATCCAGTCAAGAGATGAGGTGAAAG 293
 QY 41 ValValGlyArgGlyAlaPheGlyValValCysLysAlaLysTyrPArgAlaLysAspVal 60
 Db 294 GTTGTGGAAAGAGAGCCTT-GGAGTTGTTTGCAAGCTAAGTGAGGACAAAGATGTT 352
 QY 61 AlaIleuysGlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
 Db 353 GCTATTAAACAAATGAAAGATGGAATCTGAGGAAAGCGTTATTATGAGCTTCGCGAG 412
 QY 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnProVal 100
 Db 413 TTATTCCTCGTGAACCACTCATATTGTAAAGCTTTATGAGAGCTGTTGAATCCAGTG 472
 QY 101 CysLeuValMetGluTyrAlaGluGlyLysLeu-TyrAsnValLeuHisGlyAlaGlu 120
 Db 473 TGTCTTGATGAAATATGCTGAAGGGGGCTCTTACTATATAGTGCTGATGGTGCTGA 532
 QY 120 uProleuProtyrTyrThrAlaAlaHisAlaMetSerTyrCysLeuGlnCysSerGlnGlu 140
 Db 533 ACCATTGCAATATTAATATGCTGGCCAGCAATAGTGGTTCACATGTTTCCCAAG 592
 QY 140 ValAlaIleTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysProPr 160
 Db 593 AGTGCTTATCTTCACAGCAATGCAACCAAGCCCTAATTACAGGGACCTGAAACACC 652
 QY 160 AsnLeuLeuLeuValAlaGlyGlyThrValLeu-LysIleCysAspPheGlyThrAlaC 180
 Db 653 AAACCTACTGCTGGTGGAGGGGAGCAGTCTTAACAAACTTGATTCGTAACAGCT 712
 QY 180 ysaaplleGlnThrHisMetThrAsnAsnLysGly 191
 Db 713 GTGACATTGCAGACACATGACCAATTAACCAAGGC 747
 RESULT 14
 LOCUS AW960377
 DEFINITION ESTJ372448 MAGE resequences, MAGF Homo sapiens cDNA, mRNA sequence.
 ACCESSION AW960377
 VERSION AW960377.1 GI:8150061
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 542)
 AUTHORS Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaepard,R., Gay,C., Holt
 I.B., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
 Quakenbush,J.
 TITLE Assessment of gene expression patterns in a model of colon tumor
 metastasis using a 19,200 element cDNA microarray
 JOURNAL Unpublished (2000)
 COMMENT Contact: John Quackenbush
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 3528
 Fax: 301 838 0208
 Email: johnq@tigr.org
 Plate: 145
 Seq primer: Reverse.
 FEATURES
 Location/Qualifiers
 1..542
 /organism="Homo sapiens"
 /db xref="taxon:9606"

Qy 192 rAlaAlaTP-MetAlaProGluValPheGluGlySerAsnTyrSerGlu 208
|||
:::
Db 772 TGCTGCTTGGAATTGGCCCTGAAAGTTTCGACAGGTAAATATTACCGTGAA 821
|||
:::
|||

Search completed: December 10, 2002, 03:43:00
Job time : 2211 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 10, 2002, 02:07:24 ; Search time 71 seconds
(without alignments)

1308.775 Million cell updates/sec

Title: US-09-830-144-2_COPY_1_303

Perfect score: 1615

Sequence: 1 MSTASASSSSSSSSAGEMIE.....MTHLMRYFGADPELPQVPCQ 303

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 1533381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09830144/runat_04122002_141511_2942/app_query.fasta_1.455
-DB=Issued Patents NA -QMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09830144@cgn2_1_31@runat_04122002_141511_2942 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DBEXT=7

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/5A COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTCUT COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1615	100.0	1788	US-09-529-279-14	Sequence 14, Appl
2	1615	100.0	2656	US-08-685-625A-5	Sequence 5, Appl
3	1615	100.0	2856	US-09-529-279-3	Sequence 3, Appl
4	1604	99.3	2443	US-08-685-625A-1	Sequence 1, Appl
5	442	27.4	2120	US-09-221-235-4	Sequence 4, Appl
6	442	27.4	2120	US-09-221-928-4	Sequence 4, Appl
7	442	27.4	2120	US-09-221-527-4	Sequence 4, Appl
8	442	27.4	2120	US-09-221-236-4	Sequence 4, Appl
9	442	27.4	2120	US-09-221-416-4	Sequence 4, Appl
10	442	27.4	2120	US-09-221-245-4	Sequence 4, Appl
11	442	27.4	2120	US-09-163-115-4	Sequence 4, Appl
12	442	27.4	2120	US-09-221-528-4	Sequence 4, Appl

13	442	27.4	2120	4	US-09-593-553-4	Sequence 4, Appl
14	442	27.4	2120	4	US-09-221-237-4	Sequence 4, Appl
15	433	26.8	1365	3	US-09-221-235-6	Sequence 6, Appl
16	433	26.8	1365	3	US-09-221-928-6	Sequence 6, Appl
17	433	26.8	1365	3	US-09-221-527-6	Sequence 6, Appl
18	433	26.8	1365	3	US-09-221-236-6	Sequence 6, Appl
19	433	26.8	1365	3	US-09-221-416-6	Sequence 6, Appl
20	433	26.8	1365	4	US-09-221-245-6	Sequence 6, Appl
21	433	26.8	1365	4	US-09-163-115-6	Sequence 6, Appl
22	433	26.8	1365	4	US-09-221-528-6	Sequence 6, Appl
23	433	26.8	1365	4	US-09-593-553-6	Sequence 6, Appl
24	433	26.8	1365	4	US-09-221-237-6	Sequence 6, Appl
25	410	25.4	2505	4	US-09-291-839-3	Sequence 3, Appl
26	410	25.4	3025	4	US-09-291-839-1	Sequence 1, Appl
27	401	24.8	3389	1	US-08-395-580-1	Sequence 1, Appl
28	401	24.8	3426	1	US-08-205-018-1	Sequence 1, Appl
29	398	24.6	2890	1	US-07-928-464-1	Sequence 1, Appl
30	398	24.6	2890	5	PCT-US93-07347-1	Sequence 1, Appl
31	398	24.6	3033	1	US-08-003-311B-1	Sequence 1, Appl
32	398	24.6	3033	1	US-08-261-432-1	Sequence 1, Appl
33	357.5	22.1	2574	4	US-09-142-529-2	Sequence 2, Appl
34	352	21.8	4508	5	PCT-US93-06251-34	Sequence 34, Appl
35	350	21.7	3132	1	US-08-167-919A-9	Sequence 9, Appl
36	350	21.7	3132	3	US-08-715-106-9	Sequence 15, Appl
37	344.5	21.3	3254	1	US-08-162-809-15	Sequence 257, App
38	339.5	21.0	3516	3	US-09-188-930-257	Sequence 1, Appl
39	338.5	21.0	4304	4	US-08-368-776A-1	Sequence 1, Appl
40	338.5	21.0	4304	5	PCT-US96-00419-1	Sequence 1, Appl
41	333.5	20.7	4529	2	US-08-449-645A-16	Sequence 16, Appl
42	333.5	20.7	4529	2	US-08-702-367A-16	Sequence 16, Appl
43	333.5	20.7	4529	5	PCT-US95-04681-16	Sequence 16, Appl
44	333	20.6	1398	2	US-08-604-989A-9	Sequence 9, Appl
45	333	20.6	1521	2	US-08-604-989A-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1

US-09-529-279-14
; Sequence 14, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529, 279
; CURRENT FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 14
; LENGTH: 1788
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)..(1776)
US-09-529-279-14

Alignment Scores:
Pred. No.: 9.7e-187
Score: 1615.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 4
Length: 1788
Matches: 303
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0

US-09-830-144-2_COPY_1_303 (1-303) x US-09-529-279-14 (1-1788)

Db 483 TGTCTGTGTGAATATGCTGAAGGGGGCTCTTTATATAATGTGTGTCATGGTCTGAA 542
QY 121 ProLeuProTyrThrAlaAlaHisAlaMetSerTyrCysLeuGlnCysSerGlnGly 140
Db 543 CCATTGCCATATTACTGCTGCCACGCAATGAGTTGGTGTACAGTGTCCCAAGGA 602
QY 141 ValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysPro 160
Db 603 GTGGCTTATCTTACAGCATGCAACCAAGCGCTAATTCACAGGAGACCTGAACACCA 662
QY 161 AsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCys 180
Db 663 AACTTACTGTGTGGTGGAGGGGACAGTTCTAAATAATTTGGTACAGCCTGT 722
QY 181 AspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyrMetAlaProGluVal 200
Db 723 GACATTTCAGACACATGACCATACACAGGGAGTGTCTTGGATGGCACCTGAAGTT 782
QY 201 PheGluGlySerAsnTyrSerGluLysCysAspValPheSerTyrGlyIleLeuTyr 220
Db 783 TTTGAAGGTAGTAATACAGTGAATAATGTACCTCTTCAGCTGGGGTATTATTCTTGG 842
QY 221 GluValIleThrArgLysProPheAspGluIleGlyGlyProIleLeuMet 240
Db 843 GAAGTGATACGGCTCGGAACCCCTTTGATGAGATTGGTGGCCACGCTTCCGAATCATG 902
QY 241 TrpAlaValHisAsnGlyThrArgProLeuIleLysAsnLeuProLysProIleGlu 260
Db 903 TGGGCTGTTCAATATGGTACTCCACACCACTGATTAATAATTTACTAAGCCCATGAG 962
QY 261 SerLeuMetThrArgCysTyrSerLysAspProSerGlnArgProSerMetGluGluIle 280
Db 963 AGCCTGATGACTCGTTGGTCTAAGATCCCTCCAGCGCCCTTCAATGAGGAAAT 1022
QY 281 ValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyr 300
Db 1023 GTGAAATAATGACTCACTTGTATGCGGTACTTTCCAGGAGCAGATGAGCCATTACAGTAT 1082
QY 301 ProCysGln 303
Db 1083 CCTTGTGAC 1091

RESULT 3

US-09-529-279-3
; Sequence 3, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529,279
; CURRENT FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2656
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (183)..(1919)
US-09-529-279-3
Alignment Scores:
Pred. No.: 1.83e-186 Length: 2656
Score: 1615.00 Matches: 303
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0
US-09-830-144-2_COPY_1_303 (1-303) x US-09-529-279-3 (1-2656)
QY 1 MetSerThrAlaSerAlaAlaSerSerSerSerSerSerAlaGlyGluMetIleGlu 20
Db 183 ATGTCTACAGCCTGCGCGCTCCTCCTCCTCCTCGCTTCGCGCGGTGAGATGATCGAA 242
QY 21 AlaProSerGlnValLeuAsnPheGluGluIleAspTyrLysGluIleGluValGluGlu 40
Db 243 GCCCTTCCAGAGTCCCTCAACTTTGAAGAGATCGACTACAGGAGATCGAGGTGGAAGAG 302
QY 41 ValValGlyArgGlyAlaPheGlyValValCysLysAlaLysTyrArgAlaLysAspVal 60
Db 303 GTTCTTGAAGAGAGGAGCCTTTGGAGTTGTTGCAAGCTAAGTGGAGAGCAAAAGATGTT 362
QY 61 AlaIleLysGlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
Db 363 GCTATTAAACAAATAGAAAGTGAATCTGAGAGGAAAGCGTTTATTGTAGAGCTTCGCGCAG 422
QY 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnProVal 100
Db 423 TTATCCCGTGTGAACCATCTCTAATATTGTAAGCTTTTATGAGCCTCTCTGAAATCCAGTG 482
QY 101 CysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGlyAlaGlu 120
Db 483 TGTCTTGTGATGGAATATGCTGAAGGGGGCTCTTTATATAATGTGCTGCTGCTGAA 542
QY 121 ProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeuGlnCysSerGlnGly 140
Db 543 CCAITTCATATTACTGCTGCCACGCAATGAGTTGGTGTGTACAGTGTTCCTCAAGGA 602
QY 141 ValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysPro 160
Db 603 GTGGCTTATCTTTCAGCATGCAACCAAGCGCTAATTCACAGGAGCTGAAACACCA 662
QY 161 AsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCys 180
Db 663 AACTTACTGTGTGTTCAGGGGGGACAGTTCTAAATAATTTGTGATTTTGGTACAGCCTGT 722
QY 181 AspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyrMetAlaProGluVal 200
Db 723 GACATTTCAGACACATGACCAATAACAAGGGAGTGTCTTGGATGGCACCTGAAAGTT 782
QY 201 PheGluGlySerAsnTyrSerGluLysCysAspValPheSerTyrGlyIleLeuTyr 220
Db 783 TTTGAAGGTAGTAATACAGTGAATAATGTACGCTTTCAGCTGGGGTATTATTCTTTGG 842
QY 221 GluValIleThrArgAtqLysProPheAspGluIleGlyGlyProAlaPheArgIleMet 240
Db 843 GAAGTGATAACGCTCGGAACCCCTTTGATGAGATTGGTGGCCACGCTTCCGAAATCATG 902
QY 241 TrpAlaValHisAsnGlyThrArgProLeuIleLysAsnLeuProLysProIleGlu 260
Db 903 TGGGCTGTTCAATATGGTACTCCAGCACCACTGATTAATAATTTACCTAAGCCCATGAG 962
QY 261 SerLeuMetThrArgCysTyrSerLysAspProSerGlnArgProSerMetGluGluIle 280
Db 963 AGCCTGATGACTCGTTGTGTCTAAGATCCCTTCCAGCGCCCTTCAATGAGGAAAT 1022
QY 281 ValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyr 300
Db 1023 GTGAAATAATGACTCACTTGTATGCGGTACTTTCCAGGAGCAGATGAGCCATTACAGTAT 1082
QY 301 ProCysGln 303
Db 1083 CCTTGTGAC 1091

RESULT 4
US-08-685-625A-1
; Sequence 1, Application US/08685625A

Patent No. 5945301
GENERAL INFORMATION:
APPLICANT: UENO, Naoto
APPLICANT: MATSUMOTO, Kunihiko
APPLICANT: IRIE, Kenji
TITLE OF INVENTION: NOVEL KINASE IN TGF-BETA FAMILY SIGNAL
TITLE OF INVENTION: TRANSDUCTION SYSTEM
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESS: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22113-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/685,625A
FILING DATE: 24-JUL-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-253549
FILING DATE: 29-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: Meuth, Donna M.
REGISTRATION NUMBER: 36,607
REFERENCE/DOCKET NUMBER: 001560-267
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2443 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 157..1893
US-08-685-625A-1
Alignment Scores:
Pred. No.: 3,5e-185 Length: 2443
Score: 1604.00 Matches: 301
Percent Similarity: 99.67% Conservative: 0
Best Local Similarity: 99.67% Mismatches: 1
Query Match: 99.32% Indels: 0
Gaps: 0
US-09-830-144-2_COPY_1_303 (1-303) x US-08-685-625A-1 (1-2443)
QY 1 MetSerThrAlaSerAlaAspSerSerSerSerSerAlaGluMetIleGlu 20
DB 157 ATGTGCACAGCCCTCCGCCCTCGTCCTCTCTCTCTGCGACGAGATGATCGAA 216
QY 21 AlaProSerGluValIleuAsnPhgIleuIleAspTyrLysGluIleGluValIleGlu 40
DB 217 GCGCCCTCCAGAGTCTCTGAATCTTGGAGATCGACTACAGAGAGATCGAGGTGGAAG 276
QY 41 ValIleGluIleArgGlyAlaIlePheGlyValValCysLysAlaLysTyrPheArgAlaLysAspVal 60
DB 277 GTTGTGGAAGAGGAGCTTTTGGAGTAGTTTCCAAAGCTAACTGAGAGCAAAAAGATGTC 336
QY 61 AlaIleLysGluIleGluSerGluSerGluArgLysAlaPheIleValIleGluLeuArgGln 80
DB 337 GCTATTAAACAGATGAAAGTAGTCTGAGAGAAAGGCTTTTCATTGTGAGAGCTCGGCGAG 396
QY 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnProVal 100

DB 397 TTGTGCGGTGTAACCACTTCAACATTGTCAAGTTGACGAGCTGCTGAAATCCAGTA 456
QY 101 CysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValIleuHisGlyAlaGlu 120
DB 457 TGTCTTGATGAAATATGACAGAGGGGCTCTTGTATATATGCTGCATGCTGCTGAA 516
QY 121 ProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeuGlnCysSerGlnGly 140
DB 517 CCATTGCTTACTACACTGCTGCTCATGCGCATGAGCTGGGTGTACAGTGTCCCAAGCA 576
QY 141 ValAlaTyrLeuHisSerMetGlnProLysAlaIleuIleHisArgAspLeuLysProPro 160
DB 577 GTGCTTACTGACACATGACAGCCCAAGCCCTGATTAACAGGAGCTCAAGCTTCCA 636
QY 161 AsnLeuLeuValAlaGlyGlyThrValIleuLysIleCysAspPheGlyThrAlaCys 180
DB 637 AACTTGCTGCTGGTTGACAGAGGAGAGCTTCTAAAATCTGCGCATTTTGTACAGCTTGT 696
QY 181 AspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyrMetAlaProGluVal 200
DB 697 GACATCCAAACACACATGACCAATTAAGGAGAGTGTGCTTGATGGCGCTGAAGTG 756
QY 201 PheGluLysSerAsnTyrSerGluLysCysAspValPheSerTyrGlyIleIleLeuTyr 220
DB 757 TTGAAAGTAGCAATTAACGTGAAGAGTGTGATCTTCACTGAGGGGTATTAATCCCTGTG 816
QY 221 GluValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMet 240
DB 817 GAATGATTAACACGCGGAAACCTTCGATGAGATGGTGGCCACTTTCAGAAATCATG 876
QY 241 TyrAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGlu 260
DB 877 TGGGCTTTCATATATGACCTTCACCAACCATGATCAAAATTTAATCAAGCCCATTTGAG 936
QY 261 SerLeuMetThrArgCysTyrSerLysAspProSerGluArgProSerMetGluGluIle 280
DB 937 AGCTTATGACAGCGCTGTGTCTTAAGACCATTCACAGGCCCTTCATATGGAGAAATT 996
QY 281 ValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyr 300
DB 997 GTGAAATATATGACTCACTTGATGGGTACTTCCAGAGGCGGATGAGCATTAAGCAT 1056
QY 301 ProCys 302
DB 1057 CCTTGT 1062
RESULT 5
US-09-221-235-4
Sequence 4, Application US/09221235
Patent No. 6043040
GENERAL INFORMATION:
APPLICANT: Acton, Susan
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
FILE REFERENCE: MNI-050
CURRENT FILING DATE: 1998-12-28
EARLIER APPLICATION NUMBER: 09/163,115
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 2120
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (47)..(1411)
US-09-221-235-4
Alignment Scores:
Pred. No.: 1,02e-43 Length: 2120
Score: 442.00 Matches: 105

```
Percent Similarity: 55.33% Conservative: 56
Best Local Similarity: 36.08% Mismatches: 98
Query Match: 27.37% Indels: 32
DB: 3 Gaps: 12

US-09-830-144-2_COPY_1_303 (1-303) x US-09-221-235-4 (1-2120)

QY 10 SerSerSerSerAlaGlyGluMetIle-----GluAlaProSerGlnVal 25
Db 2 TCGACCCACGGCTCCGGTGAAGTAAATACATTGTTCATTATGAGATGCTCTCGGT 61

QY 26 LeuAsnPheGluLeuIleAspTyrLysGluIleValGluValValGlyArgGly 45
Db 62 GCCTCCTTTGTGCAATTAATTTGATGCTGCTGAGTATTTTGAACCTCGGTGAGGA 121

QY 46 AlaPheGlyValValCysLysAlaLysTrp-----ArgAlaLysAspValAlaLys 53
Db 122 AGTTTGGGAGTGTTCATCGAGCAATGATATCACAGCAAGGAGTGTCTGTAAG 181

QY 64 -----GlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
Db 182 AAGCTCTCAAAATAGAAAGGAGGAGAA-----ATA 214

QY 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnPro--- 99
Db 215 CTCAGTGTCTCAGTCACAGAAACATCATCCAGTTTATGGAGTAAATCTTGAACCTCC 274

QY 100 ---ValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGly 118
Db 275 AACTATGGCATTTGCAGAAATATGCTTCTCTGGGATCAGTCTATGATTACATTAAACAGT 334

QY 119 -----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeuGln 136
Db 335 AACAGAGTGGAGGAGATG-----GATATGGATCACATTATGACCTGGGCCACTGAT 385

QY 137 CysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAsp 156
Db 386 GTAGCCAAAGGAATGCATTATTACATATGGAGGCTCTGTCAAGGTGATTACACAGAC 445

QY 157 LeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPhe 176
Db 446 CTCAGTCAAGAAACGTTGTTATAGTCTCTGATGGA---GTACTGAAGATCTGTGACTTT 502

QY 177 GlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys-----GlySerAlaAla 194
Db 503 GGT---GCCTCTCGGTTCCATTAACCATACACACATGTCCTTGGTTGGAACCTTCCCA 559

QY 195 TrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSer 214
Db 560 TGGATGGCTCCAGAAAGTTATCCAGAGTCTCCTGTGTCAGAAACTTGTGCACATATTCC 619

QY 215 TrpGlyIleLeuLeuTrpGluValIleThrArgArgLysProPheAspGluIleGly 234
Db 215 TrpGlyIleLeuLeuTrpGluValIleThrArgArgLysProPheAspGluIleGly 234

; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,928
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2120
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (47)..(1411)
US-09-221-928-4

Alignment Scores: 1.02e-43 Length: 2120
Pred. No.: 442.00 Matches: 105
Score: 442.00 Conservative: 56
Percent Similarity: 55.33% Mismatches: 98
Best Local Similarity: 36.08% Indels: 32
Query Match: 27.37% Gaps: 12
DB: 3

US-09-830-144-2_COPY_1_303 (1-303) x US-09-221-928-4 (1-2120)

QY 10 SerSerSerSerAlaGlyGluMetIle-----GluAlaProSerGlnVal 25
Db 2 TCGACCCACGGCTCCGGTGAAGTAAATACATTGTTCATTATGAGATGCTCTCGGT 61

QY 26 LeuAsnPheGluLeuIleAspTyrLysGluIleValGluValValGlyArgGly 45
Db 62 GCCTCCTTTGTGCAATTAATTTGATGCTGCTGAGTATTTTGAACCTCGGTGAGGA 121

QY 46 AlaPheGlyValValCysLysAlaLysTrp-----ArgAlaLysAspValAlaLys 53
Db 122 AGTTTGGGAGTGTTCATCGAGCAATGATATCACAGCAAGGAGTGTCTGTAAG 181

QY 64 -----GlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
Db 182 AAGCTCTCAAAATAGAAAGGAGGAGAA-----ATA 214

QY 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnPro--- 99
Db 215 CTCAGTGTCTCAGTCACAGAAACATCATCCAGTTTATGGAGTAAATCTTGAACCTCC 274

QY 100 ---ValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGly 118
Db 275 AACTATGGCATTTGCAGAAATATGCTTCTCTGGGATCAGTCTATGATTACATTAAACAGT 334

QY 119 -----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeuGln 136
Db 335 AACAGAGTGGAGGAGATG-----GATATGGATCACATTATGACCTGGGCCACTGAT 385

QY 137 CysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAsp 156
Db 386 GTAGCCAAAGGAATGCATTATTACATATGGAGGCTCTGTCAAGGTGATTACACAGAC 445

QY 157 LeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPhe 176
Db 446 CTCAGTCAAGAAACGTTGTTATAGTCTCTGATGGA---GTACTGAAGATCTGTGACTTT 502

QY 177 GlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys-----GlySerAlaAla 194
Db 503 GGT---GCCTCTCGGTTCCATTAACCATACACACATGTCCTTGGTTGGAACCTTCCCA 559

QY 195 TrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSer 214
Db 560 TGGATGGCTCCAGAAAGTTATCCAGAGTCTCCTGTGTCAGAAACTTGTGCACATATTCC 619

QY 215 TrpGlyIleLeuLeuTrpGluValIleThrArgArgLysProPheAspGluIleGly 234
Db 620 TATGGTGTGTTCTCTGGGAGATGCTAACCAAGGAGGATCCCTTTTAAAGGTTTGGAAAGA 679

QY 235 ProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgProProLeuIleLys 253
Db 680 -----TTACAAGTAGCTTGGCTTGTAGTGAAAMAAACAGAGATTAACCATTTCCAGC 733

QY 254 AsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGln 273
Db 734 AGTTGCCCCAGAAAGTTTGTGTAACCTGTTACATCAGTGTGGGAAGCTGATGCCAAGAAA 793

QY 274 ArgProSerMetGluGluIleValLysIleMet 284
Db 794 CGGCCATCAITCAAGCAAAATCAITTCATCTG 826

RESULT 6
US-09-221-928-4
; Sequence 4, Application US/09221928
; Patent No. 6121030
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
```

Db	620	TATGGTGTGGTTCTCTGGAGATGCTAACAGAGGAGGTCCTTTAAAGTTTGAAGA	679
Qy	235	ProAlaPheArgIleMetTyr--AlaValHisangIythrArgProProIleuIleLys	253
Db	680	-----TTACAAGTGTGGCTTGTAAGTGAAAAAACAAGAGATTAAACCATTCCAAGC	733
Qy	254	AsnIeuProLysProIleGluSerIleuMetThrArgCysTrpSerLysAspProSerGln	273
Db	734	AGTTGCCCCACAAAGTTTGTCTGAACCTGTATCACTCAGTGTGGGAAGCTGATGCCAAGAA	793
Qy	274	ArgProSerMetGluGluIleValLysIleMet	284
Db	794	CGGCATCATTTCAAGCAAAATCATTTCAATCTCTG	826

```

RESULT 7
US-09-221-527-4
Sequence 4, Application US/09221527
Patent No. 6146832
GENERAL INFORMATION:
APPLICANT: Acton, Susan
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
FILE REFERENCE: NMI-050
CURRENT APPLICATION NUMBER: US/09/221,527
CURRENT FILING DATE: 1998-12-28
EARLIER APPLICATION NUMBER: 09/163,115
EARLIER FILING DATE:
NUMBER OF SEQ ID NOS: 15
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 4
LENGTH: 2120
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (47)..(1411)
US-09-221-527-4

```

Alignment Scores:	
Pred. No.:	1.02e-43
Score:	442.00
Percent Similarity:	55.33%
Best Local Similarity:	36.08%
Query Match:	27.37%
AB:	3
	Gaps:
	12
	2120
Length:	
Matches:	105
Conservative:	56
Mismatches:	98
Indels:	32
Gaps:	12

```

US-09-830-14-2_COPY_1_303 (1-303) x US-09-221-527-4 (1-2120)

QY 10 SerSerSerSerSeraIaGlyGluMetIle-----GluAlaProSergInVal 25
      |||:::  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
Db 2 TCGACCACGCGCTCCGGTGGAAATATATCTTGTCTATTATGAGATGTCGTCTCCGGT 61
      |||:::  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
QY 26 LeuAsnHeGluGluIleAspTyrLysGluIleGluValGluGluValaIaGlyArgGly 45
      |||:::  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
Db 62 GCCCTCCTTTGGCAAAATTAATTGATGACTTCGACGTTTTTGGAAACCTCCGGTGGAGA 121
      |||:::  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
QY 46 AlaPheGlyValValaCysLysAlaIaLysTrp-----ArgAlaLysAspValaIaIleLys 63
      |||:::  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
Db 122 AGTTTGGAGCTGTTTATTCGAGCCAAATGATATCATCAGACGAAAGGAGCGCTGTAAAG 181
      |||:::  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
QY 64 -----GlnIleGluSerGluSerGluValArgLysAlaPheIleValGluLeuArgIn 80
      |||:::  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
Db 182 AAGCTCTCTCAAAATAGAGAAAGAGGCGGAA-----ATA 214
      |||:::  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
QY 81 LeuSerArgValaSnHisProAsnIleValaLysLeuTyrGlyAlaCysLeuAsnPro--- 99
      |||:::  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
Db 215 CTCAGTGTCTCCACAGTCACAGAAACATATCATCGATTTATGAGTAATCTTGAACCTCCC 214
      |||:::  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
QY 100 ---ValaCysLeuValaMetGluTyrAlaGluGlyGlySerLeuTyrAsnValaLeuHisGly 118
      |||:::  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
Db 275 AACTATGGCATTTGTCCACAGAAATATGCTTCTCTGGGATCAGCTATGATTCATTACAG 334
      |||:::  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||
QY 119 -----AlaGluProLeuProTyrTyrThrAlaAlaHisIaMetSerTrpCysLeuGln 136
      |||:::  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||  ::|||

```

Db	335	AACGAAAGTAGGAGGATG-----GATATGATGATCACTTTATGACCTGGCCACTGAT	383
Qy	137	CysSerGlnGlyValAlaIYrLeuHisSerMetGlnProLysAlaLeuIleHisArgAsp	156
Db	386	GTACCCAAAGGAAGCACTTTATTACAATGAGAGCTCTCTCAAGATGATTCACAGAGAC	445
Qy	157	LeuLysPProPheAlaLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPhe	176
Db	446	CTCAAGTCAGAAACGTTGTTATATGCTGCTGATGGA---GTACTGAAGATCTGTACTTT	502
Qy	177	GlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys-----GlySerAlaAla	194
Db	503	GGT---ACCTCTCGGTTCCATTAACCAATACACACACTGTCTTGGTGGAACTTTCCCA	559
Qy	195	TryMetAlaProGluValPheGlnGlySerAsnTyrSerGlnLysCysAspValPheSer	214
Db	560	TGGATGGCTCCAGAAAGTTATCCACAGAGCTCCCTGTGTCAAAACTGTGTACACATATCC	619
Qy	215	TryGlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGly	234
Db	620	TATGATGTGGTCTCTGGAGATGCTPAACAAGGAGAGGTCCCTTTAAAGTTTGAAGGA	679
Qy	235	ProLahPheArgIleMetTrp---AlaValHisAsnGlyThrArgProProLeuIleLys	253
Db	680	-----TTACAAGAGCTTGCGTTGATGTGAAAAAACAAGAGATTAAACATTTCCAAAC	733
Qy	254	AsnLeuProLysProIleGlnSerLeuMetThrArgCysTrpSerLysAspProSerGln	273
Db	734	AGTTGCCCCAGAAAGTTTTCGTGAACGTATTACATCAGTGAGTTGGGAACGTATGCCAAGAA	793
Qy	274	ArgProSerMetGlnGluIleValLysIleMet	284
Db	794	CGGCAATCATTTCAAGCAAAATCATTTTCAATCTTG	826

```

RESULT 8
US-09-221-236-4
; Sequence 4, Application US/09221236
; Patent No. 6146841
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSARK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMT-050
; CURRENT APPLICATION NUMBER: US/09/221,236
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ. ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2120
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (47)..(1411)
; US-09-221-236-4

```

```

Alignment Scores:
Pred. No.: 1.02e-43 Length: 2120
Score: 442.00 Matches: 105
Percent Similarity: 55.33% Conservative: 56
Best Local Similarity: 36.08% Mismatches: 98
Query Match: 27.37% Indels: 32
DB: 3 Gaps: 12

US-09-830-144-2_COPY_1_303 (1-303) x US-09-221-236-4 (1-2120)

QY 10 SerSerSerSerSerAlaGlyGluMetIle-----GluAlaProSerGlnVal 25
   ||:::  ::::  ::::  ::::  |||
Db 2 TCACCCACGACGCGCGGAAAGTATATACTTGTCTATTATGAGATGTCCTCTCCGT 61
   ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::
QY 26 LeuasnPhgIugIuIleAspTyrIysGluIleGluIvalGluIvalValGlyArgGly 45
   ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::

```

```
Db 62 GCCTCCTTTGTGCAAAATAAATTTGATGACTTGTGACGTTTTTTCAGAACTGCGGTGGAGGA 121
Qy 46 AlapheGlyValValCysLysAlaLysTrp-----ArgAlaLysAspValAlaLys 63
Db 122 AGTTTGGGAGTGTATCGAGCAATGATATCACAGCAAGGAGGTGCTGTAAAG 181
Qy 64 -----GlnIleGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
Db 182 AAGCTCTCAAAATAGAGAAAGGCAGAA-----ATA 214
Qy 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnPro---- 99
Db 215 CTCAGTGTCTCAGTCACAGAAACATCCAGTTTTTATGGAGTAATTTCTTGAACCTCC 274
Qy 100 ---ValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnValLeuHisGly 118
Db 275 AACTATGGCAATTGTACAGAAATATGCTTCTCGGATCACTCTATGATTACATTAACAGT 334
Qy 119 -----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeuGln 136
Db 335 AACAGAAAGTGAGGAGATG-----GATATGGATCACATTATGACCTGGCGCACTGAT 385
Qy 137 CysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAsp 156
Db 386 GTAGCCAAAGGAATGCATTATTTACATATATGAGGCTCCTGTCAAGGTGATTACAGAGAC 445
Qy 157 LeuLysProAsnLeuLeuValAlaGlyThrValLeuLysIleCysAspPhe 176
Db 446 CTCAAAGTCAAGAAACGTTGTATAGCTGTGATGGA---GTACTGAAGATCTGTGACTTT 502
Qy 177 GlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys-----GlySerAlaAla 194
Db 503 GGT---GCCTCTCGTTCATACCAACACACACATGCTCTGTTGGTGGAACTTTCCCA 559
Qy 195 TrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSer 214
Db 560 TGGATGGCTCCAGAGTATTCAGAGTCTCCCTGTGTGAGAACTTGTGACACATATTC 619
Qy 215 TrpGlyIleIleLeuTrpGluValIleThrArgLysProPheAspGluIleGly 234
Db 620 TATGTTGGTGTCTCTGGAGATGCTAACAGAGGAGTCCCTTTTAAAGTTTGGAAAGA 679
Qy 235 ProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgProProLeuLys 253
Db 680 -----TTACAAGTAGCTTGGCTTGTAGTGGAAAAACAGAGAGATTAAACATTCACAGC 733
Qy 254 AsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGln 273
Db 734 AGTTGCCCCAGAAAGTTTGTCTGAACCTGTACATCAGTGTGGGAAGCTGATGCCAAGAAA 793
```

RESULT 9

```
US-09-221-416-4
; Sequence 4, Application US/09221416
; Patent No. 6153417
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,416
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2120
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

```
; NAME/KEY: CDS
; LOCATION: (47)..(1411)
US-09-221-416-4
Alignment Scores: 1.02e-43 Length: 2120
Pred. No.: 442.00 Matches: 105
Score: 55.33% Conservative: 56
Best Local Similarity: 36.08% Mismatches: 98
Query Match: 27.37% Indels: 32
DB: 3 Gaps: 12
US-09-830-144-2_COPY_1_303 (1-303) x US-09-221-416-4 (1-2120)
Qy 10 SerSerSerSerAlaGlyGluMetIle-----GluAlaProSerGlnVal 25
Db 2 TCGACCCACGGCTCCGGTGGAGTATATACTTGTCTATTATGAGATGTCGTCTCTCGGT 61
Qy 26 LeuAsnPheGluGluIleAspTyrLysGluIleGluValGluGluValValGlyArgGly 45
Db 62 GCCTCCTTTGTGCAAAATAAATTTGATGACTTGCAGTTTTTTGAAAACTGCGGTGGAGGA 121
Qy 46 AlapheGlyValValCysLysAlaLysTrp-----ArgAlaLysAspValAlaLys 63
Db 122 AGTTTGGGAGTGTATTCGAGCCAAATGGATATCACAGCAAGGAGGTGCTGTAAAG 181
Qy 64 -----GlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
Db 182 AAGCTCTCAAAATAGAGAAAGGCAGAA-----ATA 214
Qy 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnPro---- 99
Db 215 CTCAGTGTCTCAGTCACAGAAACATCATCCAGTTTTTATGGAGTAATTTCTTGAACCTCC 274
Qy 100 ---ValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnValLeuHisGly 118
Db 275 AACTATGGCAATTGTACAGAAATATGCTTCTCGGATCACTCTATGATTACATTAACAGT 334
Qy 119 -----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeuGln 136
Db 335 AACAGAAAGTGAGGAGATG-----GATATGGATCACATTATGACCTGGCGCACTGAT 385
Qy 137 CysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAsp 156
Db 386 GTAGCCAAAGGAATGCATTATTTACATATATGAGGCTCCTGTCAAGGTGATTACAGAGAC 445
Qy 157 LeuLysProAsnLeuLeuValAlaGlyThrValLeuLysIleCysAspPhe 176
Db 446 CTCAAAGTCAAGAAACGTTGTATAGCTGTGATGGA---GTACTGAAGATCTGTGACTTT 502
Qy 177 GlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys-----GlySerAlaAla 194
Db 503 GGT---GCCTCTCGTTCATACCAACACACACATGCTCTGTTGGTGGAACTTTCCCA 559
Qy 195 TrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSer 214
Db 560 TGGATGGCTCCAGAGTATTCAGAGTCTCCCTGTGTGAGAACTTGTGACACATATTC 619
Qy 215 TrpGlyIleIleLeuTrpGluValIleThrArgLysProPheAspGluIleGly 234
Db 620 TATGTTGGTGTCTCTGGAGATGCTAACAGAGGAGTCCCTTTTAAAGTTTGGAAAGA 679
Qy 235 ProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgProProLeuLys 253
Db 680 -----TTACAAGTAGCTTGGCTTGTAGTGGAAAAACAGAGAGATTAAACATTCACAGC 733
Qy 254 AsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGln 273
Db 734 AGTTGCCCCAGAAAGTTTGTCTGAACCTGTACATCAGTGTGGGAAGCTGATGCCAAGAAA 793
Qy 274 ArgProSerMetGluGluIleValLysIleMet 284
Db 794 CGGCCCATTCACAGCAAAATCATTTCAATCCTG 826
```

```

RESULT 10
US-09-221-245-4
; Sequence 4, Application US/09221245
; Patent No. 6180358
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,245
; EARLIER FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: US 09/163,115
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2120
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (47)..(1411)
US-09-221-245-4

Alignment Scores:
Pred. No.: 1,02e-43 Length: 2120
Score: 442.00 Matches: 105
Percent Similarity: 55.33% Conservative: 56
Best Local Similarity: 36.08% Mismatches: 98
Query Match: 27.37% Indels: 32
DB: Gaps: 12

US-09-830-144-2_COPY_1_303 (1-303) x US-09-221-245-4 (1-2120)
QY 10 SerSerSerSerSerAlaGlyGlyMetIle-----GluAlaProSerGlnVal 25
DB 2 TCGACCCAGCGCTCGGTGAGATTAATTAATCTTGTGATTAAGATGTCCTCGCT 61
QY 26 LeuAsnProGluGluIleAspTyrLysGluIleGluValGluGluValAlaGlyArgGly 45
DB 62 GCCTCTTGTGCAATTAATTAATTTGATGCTTGCAATTTTGAAGACTGCGTGAAGA 121
QY 46 AlaPheGlyValValCysLysAlaLysTrp-----ArgAlaLysAspValAlaIleLys 63
DB 122 AGTTTGGAGTGTATTATGAGCCAAATGGATATATACAGACAAAGAGTGGCTGTAAAG 181
QY 64 -----GlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
DB 182 AAGCTCCTCAAAATAGAGAAAGAGCAGAA-----ATA 214
QY 81 LeuSerArgValAlaAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnPro--- 99
DB 215 CTCAGTGTCTCAGTCACAGAAACATCATCCAGTTTATGAGATTAATCTTGAACCTCCC 274
QY 100 ---ValCysLeuValMetGluTyrAlaGluGlyGlySerLysLeuTyrAsnValLeuHisGly 118
DB 275 AACTATGGCATTGTACAGAAATATGCTCTCTGGGATACCTTAAGATTACATTAAACAGT 334
QY 119 -----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeuGln 136
DB 335 AACACAGATGAGGAGATG-----GATATGATCATCATTAATGACCTGGGCCACATGAT 385
QY 137 CysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAsp 156
DB 386 GTAGCCAAAGAAATGCAATTATTAATATGAGAGGCTCTGCAAGAGTATTCACAGAGAC 445
QY 157 LeuLysProProAsnLeuLeuValAlaGlyGlyThrValLysLysIleCysAspPhe 176
DB 446 CTCAGGTCAAGAAAGCTGTATTAGCTCTGATGGA---GTAATGAGATTTGTGACTTTT 502
QY 177 GlyThrAlaCysAspIleGlnThrHisMetThrAsnLeuLys-----GlySerAlaAla 194
DB 503 GGT---GCCTCGGTCCATTAACCATATCAACACACATGTCCTTGTTGGAACCTTCCA 559

```

```

QY 195 TrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSer 214
DB 560 TGGATGGCTCCAGAAAGTTATTCAGAGTCCTCGTGTGCAAACTTGTGACATATTCC 619
QY 215 TrpGlyIleIleLeuTyrGluValIleThrArgLysProPheAspGluIleGlyGly 234
DB 620 TATGTTGTGTTCTCTGGGAGATGCTTAACAGGAGATGCCCTTTAAAGTTTGAAGA 679
QY 235 ProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgProProLeuIleLys 253
DB 680 -----TTACAGTAGCTTGCTGTGAGTGAAGAAAAACAGAGATTAACTTCCAAAC 733
QY 254 AsnLeuProLysProIleGluSerLeuMetThrArgCysTyrSerLysAspProSerGln 273
DB 734 AGTTGCCCCAGAAAGTTTGTGCTGAACGTTAATCATCATGTTGGAACTGATGCCAAGAA 793
QY 274 ArgProSerMetGluGluIleValLysIleMet 284
DB 794 CGGCATCATTCAGCAAAATCATTCATTCCTG 826

RESULT 11
US-09-163-115-4
; Sequence 4, Application US/09163115A
; Patent No. 6183962
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/163,115A
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2120
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (47)..(1411)
US-09-163-115-4

Alignment Scores:
Pred. No.: 1,02e-43 Length: 2120
Score: 442.00 Matches: 105
Percent Similarity: 55.33% Conservative: 56
Best Local Similarity: 36.08% Mismatches: 98
Query Match: 27.37% Indels: 32
DB: Gaps: 12

US-09-830-144-2_COPY_1_303 (1-303) x US-09-163-115-4 (1-2120)
QY 10 SerSerSerSerSerAlaGlyGlyMetIle-----GluAlaProSerGlnVal 25
DB 2 TCGACCCAGCGCTCGGTGAGATTAATTAATCTTGTGATTAAGATGTCCTCGCT 61
QY 26 LeuAsnProGluGluIleAspTyrLysGluIleGluValGluGluValAlaGlyArgGly 45
DB 62 GCCTCTTGTGCAATTAATTAATTTGATGCTTGCAATTTTGAAGACTGCGTGAAGA 121
QY 46 AlaPheGlyValValCysLysAlaLysTrp-----ArgAlaLysAspValAlaIleLys 63
DB 122 AGTTTGGAGTGTATTATGAGCCAAATGGATATATACAGACAAAGAGTGGCTGTAAAG 181
QY 64 -----GlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
DB 182 AAGCTCCTCAAAATAGAGAAAGAGCAGAA-----ATA 214
QY 81 LeuSerArgValAlaAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnPro--- 99
DB 215 CTCAGTGTCTCAGTCACAGAAACATCATCCAGTTTATGAGATTAATCTTGAACCTCCC 274
QY 100 ---ValCysLeuValMetGluTyrAlaGluGlyGlySerLysLeuTyrAsnValLeuHisGly 118

```

Db	275	AACTATGGCAATGTCACAGAAATATGCTTCTCTGGATCAGCTCTATGATTAACATTAACAGT	334	Db	2	TCGACCCACGCGTCCGGTGGAAAGTATAATCTTTGTGATTATGAGATGTCGTCTCTCGGT	61
Qy	119	-----AlaGluProLeuProTyrThrAlaAlaHisAlaMetSerTrpCysLeuGln	136	Qy	26	LeuAsnPheGluGluIleAspTyrLysGluIleGluValGluGluValValGlyArgGly	45
Db	335	AACAGAACTGAGGAGATG-----GATATGGATCACATTAATGACCTGGGCCACTGAT	385	Db	62	GCTCTCTTTGTGCAAAATTAATTTGATGACTTGGCAGCTTTTGGAAAACTCGCGTGAGGA	121
Qy	137	CysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAsp	156	Qy	46	AlaPheGlyValValCysLysAlaLysTrp-----ArgAlaLysAspValAlaIleLys	63
Db	386	GTAGCAAAAGGAATGCAATTAATTTACATATGGAGGCTCCTGTCAGAGGATTCACAGAGAC	445	Db	122	AGTTTGGAGTGTTCATCGAGCCAAATGATATCACAGGACAAAGAGGTGCTGTAAAG	181
Qy	157	LeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPhe	176	Qy	64	-----GlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln	80
Db	446	CTCAAGTCAAGAACTGTTATGATGCTCTGATGGA-----GTACTGAAGATCTGTGACTTT	502	Db	182	AAGCTCTCTCAAAATAGAGAAAGAGGCAGAA-----ATA	214
Qy	177	GlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys-----GlySerAlaAla	194	Qy	81	LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnPro----	99
Db	503	GGT---GCCTCTCGGTTCCATAACCATACACACACATGTCCTTGGTGGAACTTTCCCA	559	Db	215	CTCAGTGTCTCTAGTCACAGAAACATCATCCAGTTTATGGAGTAATTTCTTGAACCTCCC	274
Qy	195	TrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSer	214	Qy	100	---ValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGly	118
Db	560	TGGATGGCTCCAGAAAGTATCCAGAGTCTCCCTGTGTGCAGAAACTTGTGCACATATTC	619	Db	275	AACATGGCATTTGTACAGAAATATGCTTCTCTGGGATCACTCTATGATTACATTAAACAGT	334
Qy	215	TrpGlyIleIleLeuTrpGluValIleThrArgLysProPheAspGluIleGlyGly	234	Qy	119	-----AlaGluProLeuProTyrThrAlaAlaHisAlaMetSerTrpCysLeuGln	136
Db	620	TATGGTGTGGTCTCTCTGGAGATGCTTAAACAGGAGGTCCCTTTAAAGGTTTGGAAAGGA	679	Db	335	AACAGAACTGAGGAGATG-----GATATGGATCACATTAATGACCTGGGCCACTGAT	385
Qy	235	ProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgProProLeuLys	253	Qy	137	CysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAsp	156
Db	680	-----TTACAAGTAGCTTGCTTGTAGTGGAATAACAGAGAGATTAAACATTTCCAAAGC	733	Db	386	GTAGCAAAAGGAATGCAATTAATTTACATATGGAGGCTCTCTGTCAGAGTGTTCACAGAGAC	445
Qy	254	AsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGln	273	Qy	157	LeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPhe	176
Db	734	AGTTGCCCCAGAAAGTGTGTGTAAGTCTGTATCATCAGTGTGGGAAGCTGATGCCAAGAA	793	Db	446	CTCAAGTCAAGAACTGTTATGATGCTCTGATGGA-----GTACTGAAGATCTGTGACTTT	502
Qy	274	ArgProSerMetGluGluIleValLysIleMet	284	Qy	177	GlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys-----GlySerAlaAla	194
Db	794	CGGCCATCATTCAGCAAAATCATTTCAATCTCTG	826	Db	503	GGT---GCCTCTCGGTTCCATAACCATACACACACATGTCCTTGGTGGAACTTTCCCA	559
RESULT 12							
US-09-221-528-4							
; Sequence 4, Application US/09221528							
; Patent No. 6190874							
; GENERAL INFORMATION:							
; APPLICANT: Acton, Susan							
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR							
; FILE REFERENCE: MNI-050							
; CURRENT APPLICATION NUMBER: US/09/221,528							
; PRIORITY FILING DATE: 1998-12-28							
; EARLIER APPLICATION NUMBER: 09/163,115							
; PRIORITY FILING DATE: 1998-09-29							
; NUMBER OF SEQ ID NOS: 15							
; SOFTWARE: PatentIn Ver. 2.0							
; SEQ ID NO 4							
; LENGTH: 2120							
; TYPE: DNA							
; ORGANISM: Homo sapiens							
; FEATURE:							
; NAME/KEY: CDS							
; LOCATION: (47)..(1411)							
US-09-221-528-4							
Alignment Scores:							
Pred. No.:							
Score:							
Percent Similarity:							
Best Local Similarity:							
Query Match:							
DB:							
US-09-830-144-2_COPY_1_303 (1-303) x US-09-221-528-4 (1-2120)							
Qy	10	SerSerSerSerAlaGlyGluMetIle-----GluAlaProSerGlnVal	25				

446	Db	CTCAAGTCAAGAAACGGTTGTTATAGCTGCTGATGGA---GTACTGAAGATCTGTGACTTTT	502
177	QY	GlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys-----GlySerAlaAla	194
503	Db	GGT---GCCTCTCGGTTCCATACCAATACAAACACACATGCTCTGGTTGGAACTTCCCA	559
195	QY	TrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSer	214
560	Db	TGGATGGCTCCAGAAGTTATCCAGAGTCTCCCTGTGTCAGAAACTTGTGACACATATTCC	619
215	QY	TrpGlyIleLeuLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGly	234
620	Db	TATGTGTGGTTCCTCTGGGAGATGCTTAACAAGGGAGGTCCTCCCTTTAAAGGTTTGAAGA	679
235	QY	ProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgProProLeuIleLys	253
680	Db	-----TTCAAGATAGCTTGGCTTGTAGTGGGAAAAAACAAGAGAGATTACCACTTCCAAGC	733
254	QY	AsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGln	273
734	Db	AGTTGCCCAAGAGTTTGTCTGGAACCTGTATCATCAGTGTGTGGGAAGCTGATCCCAAGAA	793
274	QY	ArgProSerMetGluGluIleValLysIleMet	284
794	Db	CGGCATTCATTCAAGCAAAATCATTTCAATCCTG	826

RESUIT 15

; US-09-221-235-6
 ; Sequence 6, Application US/09221235
 ; Patent No. 6043040
 ; GENERAL INFORMATION:
 ; APPLICANT: Acton, Susan
 ; TITLE OF INVENTION: NOVEL CSAK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
 ; FILE REFERENCE: MNI-050
 ; CURRENT APPLICATION NUMBER: US/09/221,235
 ; CURRENT FILING DATE: 1998-12-28
 ; EARLIER APPLICATION NUMBER: 09/163,115
 ; EARLIER FILING DATE:
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: PatentIn Ver. 2.0

Alignment Scores:		
Pred. No.:	6.3e-43	Length:
Score:	433.00	Matches:
Percent Similarity:	56.30%	Conservative:
Best Local Similarity:	36.67%	Mismatches:
Query Match:	26.81%	Indels:
DR:	3	Gaps:
		11
		1365

US-09-830-144-2 COPY 1 303 (1-303) X US-09-221-235-6 (1-1365)

[illegible]

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 10, 2002, 02:09:35 ; Search time 98 Seconds
(without alignments)
1205.608 Million cell updates/sec

Title: US-09-830-144-2_COPY_1_303

Perfect score: 1615

Sequence: 1 MNTASASSSSSSAGEMIE.....MTHLMRYFFGADEPLQYPCQ 303

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5

Ygapop 10.0, Ygapext 0.5

Fgapop 6.0, Fgapext 7.0

Delop 6.0, Delext 7.0

Searched: 350425 seqs, 194966369 residues

Total number of hits satisfying chosen parameters: 700850

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DBV=xlp
-Q=/cgn2_1/USPTO.spool/US09830144/runat_04122002.141512.2975/app.query.fasta_1.455
-DB=Published Applications NA -QWTF=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOPEXT=0 -UNING=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09830144 @CGN 1.1.36 @runat_04122002.141512.2975
-NCPV=6 -ICPV=3 -NO_XLPX -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DSV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

```
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PTC_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PTCUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1615	100.0	1788	9 US-10-158-895-14	Sequence 14, Appl
2	1615	100.0	2656	9 US-10-158-895-3	Sequence 3, Appli
3	482.5	29.9	3454	10 US-09-969-347-226	Sequence 226, App
4	442	27.4	2120	10 US-09-757-982-4	Sequence 4, Appli

5	433	26.8	1365	10	US-09-757-982-6	Sequence 6, Appli
6	432	26.7	3111	12	US-10-014-882-1	Sequence 1, Appli
7	432	26.7	3518	12	US-10-014-882-3	Sequence 3, Appli
8	411.5	25.5	1662	9	US-09-938-842A-1014	Sequence 1014, Ap
9	410	25.4	2505	10	US-09-947-199-3	Sequence 3, Appli
10	410	25.4	2505	10	US-09-947-199-9	Sequence 9, Appli
11	410	25.4	3025	10	US-09-947-199-1	Sequence 1, Appli
12	410	25.4	3026	10	US-09-947-199-7	Sequence 7, Appli
13	392	24.3	3286	10	US-09-904-389-1	Sequence 1, Appli
14	391.5	24.2	2211	9	US-09-938-842A-1577	Sequence 1577, Ap
15	387	24.0	1428	9	US-09-938-842A-882	Sequence 882, App
16	387	24.0	2892	9	US-09-938-842A-1073	Sequence 1073, Ap
17	369	22.8	1627	10	US-09-828-313-16	Sequence 16, Appl
18	365.5	22.6	3663	10	US-09-919-172-84	Sequence 84, Appl
19	358.5	22.2	6383	9	US-09-954-531-405	Sequence 405, App
20	357.5	22.1	2574	10	US-09-735-103-2	Sequence 2, Appli
21	357.5	22.1	2574	12	US-10-045-428A-2	Sequence 2, Appli
22	351.5	21.8	1638	9	US-09-938-842A-903	Sequence 903, App
23	345.5	21.4	2661	9	US-09-938-842A-1085	Sequence 1085, Ap
24	340.5	21.1	2719	10	US-09-917-800A-1405	Sequence 1405, Ap
25	333	20.6	2000	10	US-09-977-269-1	Sequence 1, Appli
26	329.5	20.4	4165	8	US-09-969-347-207	Sequence 207, App
27	329.5	20.4	4165	8	US-08-578-684-1	Sequence 1, Appli
28	329	20.4	2456	9	US-10-186-399-1	Sequence 1, Appli
29	329	20.4	2500	10	US-09-977-269-3	Sequence 3, Appli
30	329	20.4	3454	12	US-10-044-090-48	Sequence 48, Appl
31	329	20.4	3726	10	US-09-925-302-271	Sequence 271, App
32	328.5	20.3	1092	9	US-09-938-842A-553	Sequence 553, App
33	327.5	20.3	3012	9	US-09-938-842A-1479	Sequence 1479, Ap
34	321.5	19.9	1215	10	US-09-880-107-2254	Sequence 2254, Ap
35	320	19.8	3791	10	US-09-757-100B-1	Sequence 1, Appli
36	319.5	19.8	3945	10	US-09-921-771-4	Sequence 4, Appli
37	319.5	19.8	3969	10	US-09-982-610-23	Sequence 23, Appl
38	318.5	19.7	1911	10	US-09-917-800A-1611	Sequence 1611, Ap
39	318.5	19.7	3107	10	US-09-954-456-1128	Sequence 1128, Ap
40	318.5	19.7	3348	10	US-09-982-610-34	Sequence 34, Appl
41	313.5	19.4	4698	10	US-09-895-652-1	Sequence 1, Appli
42	312.5	19.3	3393	10	US-09-922-138-3	Sequence 3, Appli
43	312.5	19.3	3633	10	US-09-922-138-1	Sequence 1, Appli
44	311.5	19.3	2015	10	US-09-954-456-1983	Sequence 1983, Ap
45	308.5	19.1	2442	10	US-09-827-949-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-10-158-895-14
; Sequence 14, Application US/10158895
; Patent No. US20020155624A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: TSUCHIYA, MASAYUKI
; APPLICANT: OHTOMO, TOSHIHIKO
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; PRIOR FILING DATE: 2002-06-03
; PRIOR FILING DATE: 2000-04-11
; PRIOR FILING DATE: 1998-10-22
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 14
; LENGTH: 1788
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)..(1776)
US-10-158-895-14

Alignment Scores:

Pred. No.:	7,64e-181	Length:	1788
Score:	1615.00	Matches:	303
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-830-144-2_COPY_1_303 (1-303) x US-10-158-895-14 (1-1788)

```
QY 1 MetSerThrAlaSerAlaAlaSerSerSerSerSerSerAlaGluMetIleGlu 20
DB 7 ATGCTACAGCCTCTCGCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCA 66
QY 21 AlaProSerGluValIleuAsnPhgIuIleAspTyrLysGluIleGluValGluGlu 40
DB 67 GCCCTTCCAGGTCCTCACTTTGAAGAGATCGACTACAGAGAGATCGAGGTGGAAG 126
QY 41 ValValGlyArgGlyAlaPheGlyValValCysLysAlaLysTyrArgAlaLysAspVal 60
DB 127 GTTGTGGAGAAGAGAGCCTTTGGAGTTGTTGCCAAAGCTAAAGTGAGAGCAAAAGATGTT 186
QY 61 AlaIleLysGlnIleGluSerGluSerGluArgLysAlaPheIleValGluuLeuArgGln 80
DB 187 GCTATTAAACAAATAGAAAGATCTGAGAGAAAGCCTTTATTGTAGAGCTTCGGCAG 246
QY 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnProVal 100
DB 247 TTATCCCGTGTGAACCATCTTAATATTGTAAGCTTTATGAGCCTCGCTGAATCAGTG 306
QY 101 CysLeuValMetGluTyrAlaGluGlyLysSerLeuTyrAsnValIleuHisGlyAlaGlu 120
DB 307 TGTCTTGATGGAATATCTGAGAGGGGCTCTTTATATNAAGTCTGATGCTGAA 366
QY 121 ProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeuGlnCysSerGlnGly 140
DB 367 CCATTGCCATATTATATCTCTGCCACGCAATGAGTGTGTTTACAGTTCCTCCAAAGCA 426
QY 141 ValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysProPro 160
DB 427 GTGGCTTACTTCTACAGCATGCAACCCAAAGCGCTTAATTCACAGGAGCCTGAACACACA 486
QY 161 AsnLeuLeuLeuValAlaGlyLysThrValLysLysIleCysAspPheGlyThrAlaCys 180
DB 487 AACTTACTGCTGCTGACAGGGGAGCAGTTCTAAAAATTGTGATTTTGTGACAGCTGT 546
QY 181 AspIleGlnThrHisMetThrAsnLeuLysGlySerAlaAlaTyrMetAlaProGluVal 200
DB 547 GACATTTCAGACACACATGACCAATTAACAAGGGAGTGTGCTTGGATGGACACCTGAAGTT 606
QY 201 PheGluGlySerAsnTyrSerGlnLysCysAspValPheSerTyrGlyIleIleLeuTyr 220
DB 607 TTTGAAGGTAGTAAATTAACGTGAATAATGTGAGCTTCAAGCTGGGATTTATTTCTTGG 666
QY 221 GluValIleThrArgArgLysProPheAspGluIleGlyLysProAlaPheArgIleMet 240
DB 667 GAAGTGATTAACGGCTCGAAGAAACCTTTGATGAGATGAGTGGCCAGCTTCCGAATCAAG 726
QY 241 TyrAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGlu 260
DB 727 TGGGCTGTTCATATAGTACTCGACCAACCATATTAATAATTTACTTAAGCCCATTTGAG 786
QY 261 SerLeuMetThrArgCysTyrPserLysAspProSerGlnArgProSerMetGluGluIle 280
DB 787 AGCTGATGATCGTCTTCTGTTGTTAAAGATCCTTCCAGGCGCTTCAATGAGGAATTT 846
QY 281 ValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyr 300
DB 847 GTGAATAATATGACTCACTTGATGCGGTACTTCCAGAGACAGATGAGCCATTACAGTAT 906
QY 301 ProCysGln 303
|||||
```

DB 907 CCTTGTACAG 915

RESULT 2

```
US-10-158-895-3
; Sequence 3, Application US/10158895
; Patent No. US20020155624A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIO
; APPLICANT: TSUCHIDA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2656
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (183)..(1919)
US-10-158-895-3
```

Alignment Scores:

Pred. No.:	1,39e-180	Length:	2656
Score:	1615.00	Matches:	303
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-830-144-2_COPY_1_303 (1-303) x US-10-158-895-3 (1-2656)

```
QY 1 MetSerThrAlaSerAlaAlaSerSerSerSerSerSerAlaGluMetIleGlu 20
DB 183 ATGCTACAGCCTCTCGCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCA 242
QY 21 AlaProSerGlnValIleuAsnPhgIuIleAspTyrLysGluIleGluValGluGlu 40
DB 243 GCCCTTCCAGGTCCTCAACTTTGAAGAGATCGACTACAGAGAGATCGAGGTGGAAG 302
QY 41 ValValGlyArgGlyAlaPheGlyValValCysLysAlaLysTyrArgAlaLysAspVal 60
DB 307 TGTGTGAAGAGAGGCTTTGGAGTTGTTGCCAAAGCTTAAGTGAAGCAAAAGATGTT 362
QY 61 AlaIleLysGlnIleGluSerGluSerGluArgLysAlaPheIleValGluuLeuArgGln 80
DB 363 GCTATTAAACAAATAGAAAGATGATCTGAGAGAAAGCCTTTATTGTAGAGCTTCGGCAG 422
QY 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnProVal 100
DB 423 TTATCCCGTGTGAACCATCTTAATATTGTAAGCTTTATGAGAGCTGTGAATCCAGTG 482
QY 101 CysLeuValMetGluTyrAlaGluGlyLysSerLeuTyrAsnValLysHisGlyAlaGlu 120
DB 483 TGTCTTGATGAGATATGCTGAAGGGGCTCTTTATATATGCTGATGCTGAA 542
QY 121 ProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeuGlnCysSerGlnGly 140
DB 543 CCATTGCCATATTATACGCTGCCACGCAATGAGTTGGTGTATTAACAGTGTCCAAAGCA 602
QY 141 ValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysProPro 160
DB 603 GTGGCTTACTTCTACAGATGCAACCAAGCGCTTAATTCACAGGAGCCTGAACACACA 662
|||||
```

```
Qy 161 AsnLeuLeuValAlaGlyThrValLeuLysIleCysAspPheGlyThrAlaCys 180
Db 663 AACTTACTGCTGCTTGCAGGGGGACAGTTCTAAATAATTTGTGTTACAGCCTGT 722
Qy 181 AspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGluVal 200
Db 723 GACATTACAGACACATGACCAATAACAAGGGAGTGCTTTGGATGGCACCCTGAAGTT 782
Qy 201 PheGluGlySerAsnTrpSerGluLysCysAspValPheSerTrpGlyIleIleLeuTrp 220
Db 783 TTTGAAGGTAGTAATTTACAGTGAATAATGTGAGCTCTTCAGCTGGGGTATTATTCTTTGG 842
Qy 221 GluValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMet 240
Db 843 GAAAGTATACCGCTCGGAAACCTTTGATGAGATTGGTGGCCAGCTTTCCGAATCATG 902
Qy 241 TrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGlu 260
Db 903 TGGGCTGTTCATAATAGTACTCCACCACTGATATAAAATTTACCTAAGGCCATTGAG 962
Qy 261 SerLeuMetThrArgCysTrpSerLysAspProSerGlnArgProSerMetGluGluIle 280
Db 963 AGCTGATGACTCGTTGGTCTAAGATCCTTCCAGCGCCCTTCAATGGAGGAATT 1022
Qy 281 ValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyr 300
Db 1023 GTGAAATAATGACTCACTTGATGCGGTACTTTCCAGGAGCAGATGAGCCATTACAGTAT 1082
Qy 301 ProCysGln 303
Db 1083 CCTTGTGCAG 1091
```

RESULT 3

```
US-09-969-347-226
; Sequence 226, Application US/09969347
; Patent No. US20020115085A1
; GENERAL INFORMATION:
; APPLICANT: Eber, Reinhard
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-69
; CURRENT APPLICATION NUMBER: US/09/969,347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/60/237,598
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: US/60/237,604
; PRIOR FILING DATE: 2000-10-03
; NUMBER OF SEQ ID NOS: 318
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 226
; LENGTH: 3454
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-347-226
```

Alignment Scores:

```
Pred. No.: 8,03e-47 Length: 3454
Score: 482.50 Matches: 107
Percent Similarity: 56.63% Conservative: 51
Best Local Similarity: 38.35% Mismatches: 98
Query Match: 29.88% Indels: 23
DB: 10 Gaps: 7
```

US-09-830-144-2_COPY_1_303 (1-303) x US-09-969-347-226 (1-3454)

```
Qy 21 AlaProSerGlnValLeuAsnPheGluGluIleAspTrpLysGluIleGluValGluGlu 40
Db 535 GCACCCGGCGCCCTCCAGCTGCCAGGATGCCCTTCCACGAGCTGCAGCTAGAGGAG 594
Qy 41 ValValGlyArgGlyAlaPheGlyValValCysLysLalLysTrpArgAlaLysAspVal 60
Db 595 ATCATCGTGTGGGGGGCTTTTGGCAAGTCTATCGGGCCCTGTGGCGTGGCGAGGAGTG 654
```

```
Qy 61 AlaIleLysGlnIleGluSerGluSerGluArgLysAlaPheIleVal----- 76
Db 655 GCAGCTAAGCGCGCGCGCTGGACCCCTGAGGAAGACCCGCGAGTCAGACGCGGAGCGTG 714
Qy 77 -----GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGly 94
Db 715 TGCAGGAAGCCCGGCTCTTTGGAGCCCTGCAGCACCACCAATAATTGCGCTTAGGGGC 774
Qy 95 AlaCysLeuAsnPro-----ValCysLeuValMetGluTyrAlaGluGlyGlySerLeu 112
Db 775 GCCTGCTCAACCCCGCCACACCTCTGCTAGTGTAGTATGAGTATGCCCGGGGTGGTGCACTG 834
Qy 113 TyrAsnValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSer 132
Db 835 AGCAGGGTGTCTGGCAGTGGCCCGGGCATGAACCTACCTACCAATGATGCCCTGTGCCCATC 882
Qy 133 TrpCysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeu 152
Db 883 TGGGCTGTGCAAGTGGCCCGGGCATGAACCTACCTACCAATGATGCCCTGTGCCCATC 942
Qy 153 IleHisArgAspLeuLysProProAsnLeuLeuVal----- 165
Db 943 ATCCACCGGACCTCAAGTCCATCAACATCCTGATCCTGGAGGCCATCGAGAACCAAC 1002
Qy 166 AlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCysAsp---IleGlnThr 184
Db 1003 CTCGCAGACACGGTGCTCAAGATCACGGAATTCGCGCCTCGCCGCGAGTGGCACAGACC 1062
Qy 185 HisMetThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGluValPheGluGlySer 204
Db 1063 ACCAAGATGAGCGCTGCGGGACCTACGCTTGATGGCGGAGGTATTCCTGCTCTCC 1122
Qy 205 AsnTyrSerGluLysCysAspValPheSerTrpGlyIleIleLeuTrpGluValIleThr 224
Db 1123 CTCCTCTCAAAGCAGTGTCTGGAGCTTCGGGGTGTGCTGTGGAGCTCTGACG 1182
Qy 225 ArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMetTrpAlaVal--- 243
Db 1183 GGGGAGGTCCCTACCGGTAGATC-----GACGCTTGGCGGTGCGGTATGGCGTGCT 1236
Qy 244 HisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGluSerLeuMet 263
Db 1237 ATGAATAAGCTGACGCTGCCCATTCCTCCACGTGCCCGAGCCCTTTGCCCGCTCTG 1296
Qy 264 ThrArgCysTrpSerLysAspProSerGlnArgProSerMetGluGluIleValLys 282
Db 1297 GAGGATGCTGGAGCCACAGACCCACCGGGCGGCAGATTTCGGTAGATCTTGAAG 1353
RESULT 4
US-09-757-982-4
; Sequence 4, Application US/09757982
; Patent No. US20020094559A1
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/757,982
; CURRENT FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 09/163,115
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2120
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (47)..(1411)
US-09-757-982-4
Alignment Scores:
Pred. No.: 2.3e-42 Length: 2120
```

Score:	442.00	Matches:	105
Percent Similarity:	55.33%	Conservative:	56
Best Local Similarity:	36.08%	Mismatches:	98
Query Match:	27.37%	Indels:	32
DB:	10	Gaps:	12

US-09-830-144-2_COPY_1_303 (1-303) X US-09-757-982-4 (1-2120)

Oy	10	SerserserserSeriGluGluMetIle-----GluAlaProSerGluVal	25
Db	2	TCGACCACCGCGCCGGTGGAGATATATACCTTTGTCTATTGAGATGTCTCTCGGT	61
Oy	26	LeuasnPheGluGluIleAspTyrLysGluIleGluValGluGluValValGlyArgGly	45
Db	62	GCCTCTTTGTGCAAATTAATTTGATGATCTTGACAGTTTTTTGAAAACTCGCGGTGAGAGA	121
Oy	46	AlaPheGluValValCysLysAlaLysTrp-----ArgAlaLysAspValAlaIleLys	63
Db	122	AGTTTGGGAGCTGTTATTATGACGCAAAATGGATATCATCAGGCAAGAGAGGTGGCTGAAG	181
Oy	64	-----GlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln	80
Db	182	AAGCTCTCAAAATATGAGAAAGAGCAGAA-----ATA	214
Oy	81	LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnPro--	99
Db	215	CTCATGTCTCCAGTACACAGAAACATCTATCCAGTTTATATGGAGTAATTTTGAACCTCCC	274
Oy	100	--ValCysLeuValMetGluTyrAlaGluGluLysSerLysLeuTyrAsnValLeuIleGly	118
Db	275	AACATGGCATTTGTACAGATATGCTTCTCTGGAGTACCTCTATGATTACATTAAACGT	334
Oy	119	-----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeuGln	136
Db	335	AACAGAACTGAGGAGATG-----GATATGATCATCATTAAGACTGGGCCACATGAT	385
Oy	137	CysserGlnGluValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisAspAsp	156
Db	386	GTAGCCAAAGGAATGTCATTATTTCATATATGAGAGCTCTGTCAAGATGATATCACAGAAC	445
Oy	157	LeuLysProProAsnLeuLeuLeuValAlaGluGlyThrValLeuLysIleCysAspPhe	176
Db	446	CTCAAGTCAAGAAACGTTGTTATGAGTCTGATGGA--GTACTGAAGATCTGTGACTTT	502
Oy	177	GlyThrAlaCysAspIleGlnThrHisMetThrAsnLys-----GlySerAlaAla	194
Db	503	GGT--GCCCTCGGTCATTAACCATACACACACACATGTCCTTGTTGGAACCTTCCCA	559
Oy	195	TrpMetAlaProGluValPheGluGlySerAsnTyrSerGlnLysCysAspValPheSer	214
Db	560	TGGATGGCTCCAGAAAGTTATCCAGAGTCCCTCGTGTCCAGAACTTGTACACTATTTCC	619
Oy	215	TrpGlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGly	234
Db	620	TATGCTGTGGTCTCTCTGGAGATCTTAACAAGGAGAGGCCCTTTAAAGTTTGGAAAGA	679
Oy	235	ProAlaPheArgIleMetTrp--AlaValHisAsnGlyThrArgProProLeuIleLys	253
Db	680	-----TTACAAAGTAGCTTGCTGCTGTATGTGAAAAAAACGAGACATTTAAACCATTTCCAAAC	733
Oy	254	AsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGln	273
Db	734	AGTTGCCCCAGAAAGTTTGTGGAAGCTGTATACATCAGTGTGGAAAGCTATGCCAAGAAA	793
Oy	274	ArgProSerMetGluGluIleValLysIleMet	284
Db	794	CGGCCATTCATTCACCAATATATTTCATCTCTG	826

```

? APPLICANT: Acton, Susan
? TITLE OF INVENTION: NOVEL CSAF-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
? FILE REFERENCE: WNI-050
? CURRENT APPLICATION NUMBER: US/09/757,982
? CURRENT FILING DATE: 2001-01-10
? PRIOR APPLICATION NUMBER: 09/163,115
? PRIOR FILING DATE: 1998-09-29
? NUMBER OF SEQ ID NOS: 15
? SOFTWARE: Patentin Ver. 2.0
? SEQ ID NO 6
? LENGTH: 1365
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (1)..(1365)
? US-09-757-982-6

```

Alignment Scores:	
Pred. No.:	1,36641
Score:	433.00
Percent Similarity:	56.30%
Best Local Similarity:	36.67%
Query Match:	26.81%
DB:	10
	Gaps: 11
	Length: 1365
	Matches: 99
	Conservative: 53
	Mismatches: 90
	Indels: 28

US-09-830-144-2_COPY_1_303 (1-303) x US-09-757-982-6 (1-1365)

Oy	27	Asn	Phe	Glu	Glu	Ile	Leu	Asp	Tyr	Leu	Ser	Glu	Ile	Ile	Glu	Val	Ala	Glu	Val	Ala	Gly	Arg	Gly	Ala	46			
Db	19	TCC	CTT	TGG	CAAA	TAA	TAA	ATT	TGA	TGC	ACTT	CTT	TTT	TGAAA	AA	CTG	CCG	GTG	GAG	GA	AG	---	---	---	78			
Oy	47	Phe	Gly	Val	Val	Cys	Leu	Ser	Ala	Leu	Tyr	Pro	---	---	---	Arg	Ala	Leu	Ser	Pro	Ala	Ile	Leu	---	63			
Db	79	TTT	GGA	AG	TCT	TTT	TGC	AG	CCAA	TG	AT	TGC	AC	AG	CA	GAG	CA	GAG	AG	TGC	CTT	TA	AA	GAA	138			
Oy	64	---	---	Gln	Ile	Glu	Ser	Glu	Ser	Glu	Ser	Glu	Val	Ala	Phe	Ile	Val	Glu	Leu	Val	Arg	Gln	Leu	---	81			
Db	139	CTC	CTC	AAA	AA	TAG	AA	AG	AA	AG	GG	CG	CA	---	---	---	---	---	---	---	---	---	---	---	171			
Oy	82	Ser	Ile	Gly	Ala	Ser	Pro	Ala	Ile	Val	Ile	Leu	Tyr	Gly	Ala	Cys	Leu	Ser	Pro	---	---	---	---	---	99			
Db	172	AGT	GCT	CTC	CA	GCT	CA	CA	AA	CA	TC	AT	CTC	CA	AGT	TTT	TA	GG	AG	TAA	TTT	TTG	AA	CTC	CC	CA	AC	231
Oy	100	Val	Cys	Leu	Val	Met	Glu	Tyr	Ile	Ala	Glu	Gly	Ser	Ile	Tyr	Asn	Val	Leu	Met	Ile	Ser	Gly	---	---	---	118		
Db	232	TAT	GCA	CTT	GT	CA	CA	GA	AT	TG	CTT	CTC	TG	GA	CTC	CT	AT	GAT	TAA	CTT	TA	CA	CT	TA	CA	CT	291	
Oy	119	---	---	Ala	Glu	Pro	Leu	Pro	Tyr	Tyr	Tyr	Thr	Ala	Ala	Ile	Ser	Leu	Pro	Cys	Leu	Gln	Cys	---	---	---	137		
Db	292	AGA	AGT	GGA	GGA	GAT	---	---	---	---	---	GAT	AT	GAT	CA	CTC	ATT	TG	CA	CTG	GG	CC	CA	CT	GAT	GT	342	
Oy	138	Ser	Gln	Gly	Val	Ala	Tyr	Leu	Met	Ser	Met	Gln	Pro	Leu	Ser	Ala	Leu	Ile	His	Arg	Ser	Leu	---	---	---	157		
Db	343	GCC	AA	GGA	AG	AT	CA	TAT	TTT	AT	CA	TAT	GA	GAG	CTC	CTG	TCA	AG	TG	AT	CA	CTT	CA	GAG	CA	CTC	402	
Oy	158	Lys	Pro	Pro	Cys	Leu	Leu	Leu	Val	Ala	Gly	Gly	Thr	Val	Leu	Lys	Ile	Cys	Asp	Phe	Gly	---	---	---	---	177		
Db	403	AAG	CA	AA	GA	AA	CG	TTT	TTA	TAG	CTG	CG	TGA	---	---	GTA	CT	GA	AG	AT	CTG	TA	CTT	TGG	---	459		
Oy	178	Thr	Ala	Cys	Asp	Ile	Gln	Thr	His	Ser	Met	Thr	Asn	Asn	---	---	Gly	Ser	Ala	Ile	Tyr	---	---	---	---	195		
Db	460	---	---	GCT	CTC	GG	TT	CC	CA	TAC	CA	TAC	CA	CA	CA	CA	AT	GCT	CTT	GG	TGG	AA	CTT	CC	CA	TGG	516	
Oy	196	Met	Ala	Pro	Glu	Val	Phe	Glu	Gly	Ser	Asn	Tyr	Ser	Glu	Lys	Cys	Asp	Val	Phe	Ser	Tyr	---	---	---	---	215		
Db	517	ATG	GCT	CC	CA	AG	TAT	TAC	CA	GAG	CTC	CC	CTG	TC	CA	GA	AA	CTT	GA	CA	CA	AT	TTC	CT	CT	CT	AT	576
Oy	216	Gly	Ile	Ile	Leu	Tyr	Pro	Glu	Val	Ile	Met	Arg	Arg	Lys	Pro	Phe	Asp	Ile	Ile	Gly	Pro	---	---	---	---	235		
Db	577	GGT	GTG	CTT	CTC	TGG	GA	GAG	CTC	TAA	CA	GGA	GAG	GCT	CC	CTT	TAA	GCTT	TGG	AA	GGA	---	---	---	---	633		
Oy	236	Ala	Phe	Arg	Ile	Met	Tyr	---	---	Ala	Val	His	Asn	Gly	Thr	Arg	Pro	Leu	Ile	Leu	Asn	---	---	---	---	254		


```

QY 77 ---GluLeuArgGluLeuSerArgValAsnHisProAsnIleValIleuTyGlyAla 95
Db 771 CGCGAGGCTCGGCTCTTCCGATGCTCGCGCACCCCAACATCATGAGTGGCGGCTGG 830
QY 96 CysLeuAsnPro-----ValCysLeuValMetGluTyAlaGluGlyGlySerIleuTy 113
Db 831 TCCCTGCACACACCGCACCTCTGCTGCTGAGATTGCCCGCGGAGCGCTCAAC 890
QY 114 AsnValIleuHisGlyAlaGluProLeuProTyTyThrAla----- 127
Db 891 CGAGCCCTGCGCGCTCCCAACCGCCCGGACCCGCGCGCCCGCGCGCGCGG 950
QY 128 -----AlaHisAlaMet---SerTyPcysLeuGlnCysSerGlnGlyValAla 142
Db 951 CGCGGATCCCTCCGACGTCGTGTCACACTGGCGCTGACATGACGCGGAGCTATGCTC 1010
QY 143 TyrLeuHisSerMetGlnProIySalLeuIleHisArgAspLeuIleProProAsnLeu 162
Db 1011 TACCTGCATGAGGAGGCTTCCTGCTGCCATCTGCACCGGACCTCAAGTCCAGCAACATT 1070
QY 163 LeuLeuValAlaGly-----GlyThrValIleuIySalIleCysAsp 175
Db 1071 TTGCTTACTTGAGAAAGATAGAACATGATGACATCTGCATTAATACTTGAGATTACAGAT 1130
QY 176 PheGlyThrAlaCysAsp---IleGlnThrHisMetThrAsnAsnIyGlySerAlaAla 194
Db 1131 TTTCGGTTGGCGGAGGATGAGACAGACACCAAAATAGACACAGAGGACCTATGGC 1190
QY 195 TrpMetAlaProGluValPheGluGlySerAsnTySerGlnIyCysAspAlaPheSer 214
Db 1191 TGGATGCGCCCGCAAGTGTATCAAGTCTCTGTTTTCGAAGGAAACGACATCTGGAGC 1250
QY 215 TrpGlyTyrIleLeuTyProGluValIleThrArgArgIyProPheAspGluIleGlyIy 234
Db 1251 TATGAGTGTGCTGTGGAGACTGCTCACCGGAGAGTCCCTATGGGGGACTTATGGC 1310
QY 235 ProAlaPheArgIleMetTyProAlaValHisAsnGlyThrArgProProLeuIleIyAsn 254
Db 1311 CTCGCCGTGGCTTATGAGGAGACATCAATAACTCATCTTGG---CCCATTCATCCACCC 1367
QY 255 LeuProIyProIleGluSerIleMetThrArgCysTyProSerIyAspProSerGlnArg 274
Db 1368 TGGCCCTGAGCGCTTTCGCCAAGCTCATGAAGAAATGCTGGCAACAGACCTCATATTGCT 1427
QY 275 ProSerMetGluGluIleValIySalIleMetThr 285
Db 1428 CCATCGTTTGGCTTAATTCTCGCAACAGTTGACT 1460

RESULT 8
US-09-938-842A-1014
: Sequence 1014, Application US/09938842A
: Patent No. US20020160378A1
: GENERAL INFORMATION:
: APPLICANT: Harper, Jeff
: APPLICANT: Krepes, Joel
: APPLICANT: Wang, Xun
: APPLICANT: Zhu, Tong
: TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
: TITLE OF INVENTION: SAME, AND METHODS OF USE
: FILE REFERENCE: SCLIP1300-3
: CURRENT APPLICATION NUMBER: US/09/938,842A
: PRIOR FILING DATE: 2001-08-24
: PRIOR APPLICATION NUMBER: US 60/227,866
: PRIOR FILING DATE: 2000-08-24
: PRIOR APPLICATION NUMBER: US 60/264,647
: PRIOR FILING DATE: 2001-01-16
: PRIOR APPLICATION NUMBER: US 60/300,111
: PRIOR FILING DATE: 2001-06-22
: NUMBER OF SEQ ID NOS: 5379
: SEQ ID NO 1014
: LENGTH: 1662
: TYPE: DNA

```

```

: ORGANISM: Arabidopsis thaliana
US-09-938-842A-1014

Alignment Scores:
Pred. No.: 6,29e-39 Length: 1662
Score: 411.50 Matches: 96
Percent Similarity: 52.35% Conservative: 60
Best Local Similarity: 32.21% Mismatches: 111
Query Match: 25,488 Indels: 31
DB: 9 Gaps: 9

US-09-830-144-2_COPY_1_303 (1-303) x US-09-938-842A-1014 (1-1662)

QY 19 IleGluAlaProSerGlnValIleuAsnPheGluGluIleAspTyTyGlySerGluVal 38
Db 772 GTTGAAATACCTACCGATGAGTAGMACTGATGAGTGGAAATGACATGGAAGACCTCAAAAT 831
QY 39 GluGluValAlaGlyArgGlyAlaPheGlyValValCysLeuSalAlaIyStrParGAlaIyS 58
Db 832 GAAAAAAGGTGGCATGTGATCATACGGGAACTATTAGAGGAACCTATTGTACTGAG 891
QY 59 AspValAlaIleuIyGlnIleGluSerGluSerGluArgIySalAheIleValGluLeu 78
Db 892 GAAGTACTCTATCAAAATTCACAGCTGAGCTGTTAATGCGAAATGCTACGAGAGTTT 951
QY 79 ArgGln-----LeuSerArgValAsnHisProAsnIleValIySerIyGly 94
Db 952 TCTCAGAGAAATATATATATATATAGAGAAAGTTCGGCATTAATAATTTGCTCCATTGCT 1011
QY 95 AlaCysLeuAsnPro-----ValCysLeuValMetGluTyAlaGluGlyIySerLeu 112
Db 1012 GCATGTACAGATATCAAAACCTTCGATTTGTACAGAGTTCTAGCTGGGGAGCATT 1071
QY 113 TyrAsnValIleuHisGlyAlaGluProLeuProTyTyThrAlaAlaHisAlaMetSer 132
Db 1072 TATGATTTCTTCACAAAACAAAGGGTT-----TTTAAATTCATCTTGTCTCAAA 1125
QY 133 TrpCysLeuGlnCysSerGlnIyValAlaTyLeuHisSerMetGlnProIySalAlaLeu 152
Db 1126 GTGCGACATCGACGCTCGTGAAGAAAGATATATATCTGCATGAA-----AACAAATAT 1176
QY 153 IleHisArgAspLeuIySProProAsnLeuLeuValAlaGlyIyThrValIleuIyS 172
Db 1177 ATTCATAGAGACTTAAAGACTGCTTAATCTTCTTATG---GACGAACTGAAGATTGTCAAA 1233
QY 173 IleCysAspPheGlyThrAlaCysAspIleGlnThrHis-----MetThrAsnAsn 189
Db 1234 GTTCCCGATTTTGGTGTGGC---AGAGTGCAGACTGAGTCAGGGGTATAGCAGCGGAA 1290
QY 190 IySglYserAlaAlaATPMeAlaProGluValPheGluGlySerAsnTyTySerGluIyS 209
Db 1291 ACAGGGCATATCCGATGAGTGGCTCCAGAGTCAATTGAGACAAACCTTATGATCATCAGG 1350
QY 210 CysAspValPheSerTrpGlyIleLeuTyProGluValIleThrArgArgIySProPhe 229
Db 1351 GCAGATGCTTCACGCTACGGCATTTGCTGTGGAACTTTTGACTGGGGAACCTCCATAT 1410
QY 230 AspGluIleGlyIyProAlaPheArgIleMetTyProAlaValHisAsnGlyThrArgPro 249
Db 1411 TCTTACTTG---ACTCCATGCAAGAGTGTGGCTGTCCAAAAGGAGCTTAAGCCA 1467
QY 250 ProLeuIleIySAsnLeuProIySProIleGluSerIleMetThrArgCysTyProSerIyS 269
Db 1468 AAAATTTCCAAAGGAACACACCCAAAATGATGAACTTTTGAGAAATCTGCGACAGCA 1527
QY 270 AspProSerGlnArgProSerMetGluGluIleValIySalIleMetThrHisSleuMetArg 289
Db 1528 GACCCAGCTTAAGACCAATTTTGCAGAAATCATTAAGAAATGCTTAAACCAACTAATCCGC 1587
QY 290 -----TyTrpPheProGly 293
Db 1588 GAGGTAATTGATTATATCATTTGCATTAAGATTAACATGCTGTTACTTTTCAGGC 1641

```


RESULT 9
 US-09-947-199-3
 ; Sequence 3, Application US/09947199
 ; Patent No. US20020127684A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Raju, Jeyaseelan
 ; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES
 ; FILE REFERENCE: MNI-068CP2
 ; CURRENT APPLICATION NUMBER: US/09/947,199
 ; CURRENT FILING DATE: 2001-09-05
 ; PRIOR APPLICATION NUMBER: 60/111,938
 ; PRIOR FILING DATE: 1998-12-11
 ; PRIOR APPLICATION NUMBER: 09/291,839
 ; PRIOR FILING DATE: 1999-04-14
 ; PRIOR APPLICATION NUMBER: 09/458,457
 ; PRIOR FILING DATE: 1999-12-10
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 3
 ; LENGTH: 2505
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(2505)
 US-09-947-199-3
 Alignment Scores:
 Pred. No.: 1.76e-38 Length: 2505
 Score: 410.00 Matches: 103
 Percent Similarity: 57.89% Conservative: 51
 Best Local Similarity: 38.72% Mismatches: 91
 Query Match: 25.39% Indels: 22
 DB: 9 Gaps: 9
 US-09-830-144-2_COPY_1_303 (1-303) x US-09-947-199-3 (1-2505)
 Qy 31 IleAspTyrLysGluIleValGluValValGlyArgGlyAlaPheGlyValVal 50
 Db 1372 CTTGAGCTCTCAGAAATTGAGTTCCATGAGATTATTGGCTCAGGTTCTTTGGGAAAGTA 1431
 Qy 51 CysIysAlaLysTrpArgAlaLysAspValAlaIleLysGlnIleGlu----- 66
 Db 1432 TATAAGGACGATGCAGAAATAAAATAGTGGCTATAAAACGTTATCGAGCCAAATACCTAC 1491
 Qy 67 ---SerGluSerGluArgLysAlaPheIleValGluLeuArgGlnLeuSerArgValAsn 85
 Db 1492 TGCTCCCAAGTCAGATGGATATGTTTGGCGAGAGTGTCCATTCTGCCAGCTCAAT 1551
 Qy 86 HisProAsnIleValLysLeuTyrGlyAlaCysLeuAsn-----ProValCysLeu 102
 Db 1552 CATCCCTCGTAATTCAGTTTGTGGTGCTTGTGAATGATCCAGCCAGTTTGGCCATT 1611
 Qy 103 ValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGlyAlaGluProLeu 122
 Db 1612 GTCACTCAATACATATCAGGGGGTCTCTGTTCTCCCTCTCATGA-GCAGAAGAGGAT 1670
 Qy 123 ProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeuGlnCysSerGln-GlyValAl 142
 Db 1671 TCTTGATTT-----GCAGTCTAAATTAAATTATTGCGATGATGTTGCCAAGGCATGGA 1724
 Qy 142 aTyrLeuHisSerMet---GlnProLysAlaLeuIleHisArgAspLeuLysProProAs 161
 Db 1725 GTACCTTCACACCTGACACAGCCA-----ATTATACATCGTGACTTTGAACATCACA 1778
 Qy 161 nLeuLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCysAs 181
 Db 1779 TATTCTTCTATGAGGATGGCATGCTGTG---GTGCGACATTTTGGAGAAATCAAGATT 1835
 Qy 181 pileGlnThr-----HisMetAsnAsnLysGlySerAlaIleTrpMetAl 197
 Db 1836 TCTACAGTCTCTGATGAAGCAACAATGACAAAACAACCTGGGAACCTCCGTTGGATGGC 1895

```
Dh 1519 TGGCCAGAGGTGTCTCTCTGCCAGCTCAACACCCTGCGTGGTTCAGTTTGGGCT 1578
Qy 95 AlaCysLeu-----AsnProValCysLeuValMetGluTyrAlaGluGlySer 111
Dh 1579 GCGTCCTGGATGACCCCGATGCTGATTTGGCATCTCAGTACAGTTCACAGAGGCTCC 1638
Qy 112 LeuTyrAsnValLeuHISGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMet 131
Dh 1639 CTGTTCTCCCTGCTCTCT-----GAACAGAGAGAAATTCCTGACTTGCAGTCTAAATTA 1692
Qy 132 SerTyrCysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMet--GlnProlys 150
Dh 1693 ATCATTTGGGTAGACGTGGCAGAGGCATGAGTACTCTCAGACGCTTGAACCCAGCA--- 1749
Qy 151 AlaLeuIleHisArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrVal 170
Dh 1750 ---ATCATACACCGCAGCAGCTGAACAGCCCAATATTTCTGCTATAGAGATGGCATGCT 1806
Qy 171 LeuLysIleCysAspPheGlyThrAlaCysAspIleGlnThr-----HisMet 186
Dh 1807 GTG---GTGGCAGATTTTGGAGATCAAGATTCTTCAGTCCCTGGATGAGATGAACAATG 1863
Qy 187 ThrAsnAsnLysGlySerAlaAlaTyrMetAlaProGluValPhe--GluGlySerAsn 205
Dh 1864 ACBAAGCAGCCGAGGAACTCGCGCTGATGGCCCCCTGAGGTGTTTCACAGTGCAGAGAA 1923
Qy 206 TyrSerGluLysCysAspValPheSerTyrGlyIleIleLeuTyrGluValIleThrArg 225
Dh 1924 TACACCATCAAGCGTATGCTTCACTTACTCCTGTGTCTGTGGAGAGCTCTCCTCATGGA 1983
Qy 226 ArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMetTyrAlaAlaHisAsn 245
Dh 1984 GAAATTCATTCCTCTCATCTCAAGCAGCCGCTGCAGCAGCAGATGTGCGATATCAACC 2043
Qy 246 GlyThrArgProProLeuIleLysAsnLeuProLysProIleGluSerLeuMetThrArg 265
Dh 2044 ---ATCAGACCGGCCATCGCGCTATTCATCTCCCAAGCCCATCTCATCCCTGCTGATG 2100
Qy 266 CysTyrSerLysAspSerGlnArgProSerMetGluGluIleVal 281
Dh 2101 GCGTGAATGATGATGCTCTGAGAGCAGCAGAGTTCTCTGAAGTGGT 2148

RESULT 11
US-09-947-199-1
; Sequence 1, Application US/09947199
; Patent No. US20020127684A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jayaseelan
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: MN1-068CP2
; CURRENT APPLICATION NUMBER: US/09/947,199
; PRIOR FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: 60/111,938
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/291,839
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 09/458,457
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3025
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (48) ..(2552)
; US-09-947-199-1

Alignment Scores:
Pred. No.: 2,34e-38 length: 3025
Score: 410.00 Matches: 103
```

```
Percent Similarity: 57.89% Conservative: 51
Best Local Similarity: 38.72% Mismatches: 91
Query Match: 25.39% Indels: 22
DB: 10 Gaps: 9

US-09-830-144-2_COPY_1_303 (1-303) x US-09-947-199-1 (1-3025)

Qy 31 IleAspTyrLysGluIleGluValGluGluValAlaGlyArgGlyAlaPheGlyValVal 50
Dh 1419 CTTGAGCTCTCGAATAATTGAGTTCCATGAGATTATGGCTCAGAGTTCTTTGGGAAGTA 1478
Qy 51 CysLysAlaLysTyrPAlaLysAspValAlaIleLysGlnIleGlu-----66
Dh 1479 TATTAAGACGATGCAGAAATTAATATGCTGCTATTAAGCTTATGAGCCAAATACCTTAC 1538
Qy 67 ---SerGluSerGluArgLysAlaPheIleValGluLeuArgGlnLeuSerArgValAsn 85
Dh 1539 TGCTCCAAAGTCAGATGTGATGATGTTTGGCCGAGAGGTGCCATTCCTGCGACGTCAT 1598
Qy 86 HisProAsnIleValLysLeuTyrGlyAlaCysLeuAsn-----ProValCysLeu 102
Dh 1599 CATCCCTGCGTAATTCAGTTGTGGGGCTTGCTGATGATCCAGCCAGTTTGGCATT 1658
Qy 103 ValMetGluTyrAlaGluGlySerLeuTyrAsnValLeuHISGlyAlaGluProLeu 122
Dh 1659 GTCACTCAATTCATATAGAGGGGTTCTCTGTTCTCCTCTTATAGA--GCAGAGAGAT 1717
Qy 123 ProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeuGlnCysSerGln--GlyValAl 142
Dh 1718 TCTTGATTT-----GCAGCTAAATTAATTAATTTGAGTGAATGTTGCCAAGGATGGA 1771
Qy 142 aTyrLeuHisSerMet--GlnProLysAlaLeuIleHisArgAspLeuLysProProAs 161
Dh 1772 GTACCTTCACAACTGCAGACAGCA-----ATTATACATCGTGAAGCTTGAACAGTACAA 1825
Qy 161 nLeuLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCysAs 181
Dh 1826 TATTTCTCTATAGAGATGAGGATGCTGTG---GTGGCAGATTTTGGAAATCAAGATT 1882
Qy 181 pIleGlnThr-----HisMetThrAsnAsnLysGlySerAlaAlaTyrMetAl 197
Dh 1883 TCTACAGTCTGTGTAAGAAATCTCACTGCGCAAAATTCATTCGCTCATCTCAAGCCAGCG 1942
Qy 197 aProGluValPhe--GluGlySerAsnTyrSerGluLysCysAspValPheSerTyrG 216
Dh 1943 TCCGAGGTGTTCACGAGCAGCACTCGGTACACCATCAAGAGATGCTTTCAGCTATGC 2002
Qy 216 yIleIleLeuTyrGluValIleThrArgArgLysProPheAspGluIleGlyProAl 236
Dh 2003 TCTGTGCTGTGGAAATTCCTCACTGCGCAAAATTCATTCGCTCATCTCAAGCCAGCG 2062
Qy 236 aPheArgIleMetTyrAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeuP 256
Dh 2063 TGGGCGAGCAGACATGCTTACACACC---ATCAAGACCTCCATGGCATTCATTC 2119
Qy 256 OlySerProIleGluSerLeuMetThrArgCysTyrSerLysAspProSerGlnArgPro 276
Dh 2120 CAAGCCCATATCATCTCTGCTGATACGAGGTGAACGCATGTCTGAAGAGAACCCGA 2179
Qy 276 rMetGluGluIleVal 281
Dh 2180 ATTTTCTGAAGTTGTC 2195

RESULT 12
US-09-947-199-7
; Sequence 7, Application US/09947199
; Patent No. US20020127684A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jayaseelan
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: MN1-068CP2
; CURRENT APPLICATION NUMBER: US/09/947,199
```


QY 126 ThrAlaAlaHisAlaMetSerTrpCysLeuGlnCysSerGlnGlyValAlaTrpLeuHis 145
DB 2146 GATGAACACGTCGATTAATATGCTTTGATGTGGCAAGGAAATGAACTACCTCCAC 2205
QY 146 SerMetGlnProLysAlaLeuLeuHisArgAspLeuLysProProAsnLeuLeuVal 165
DB 2206 AGACGTCATCT--CCAAATGTCATCGGATTTAAATCAACGAAATGTTAGTTGAC 2262
QY 166 AlaGlyGlyThrValLeuLysIleCysAspPheGly--ThrAlaCysAspIleGlnThr 184
DB 2263 AAGAACTATACAGTC--AAGGTTTGTGATTTGTCTCCCGTTTAAAGCAGCACA 2319
QY 185 HisMetThrAsnAsnLys-----GlySerAlaAlaTrpMetAlaProGluValPheGlu 202
DB 2320 TTTCTTTCATCCAAATCTGCAGCTGGACACCTGAAATGATGGACACCAAGACTACCC 2379
QY 203 GlySerAsnLysSerGlnLysCysAspValPheSerTrpGlyIleIleLeuTrpGluVal 222
DB 2380 GATGAACATCAATGAAAGTCAATGATTTTACAGCTTGGATGATTTTGTGGAGTTG 2439
QY 223 IleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMetTrpAla 242
DB 2440 GCACCTTTCGCAACGACGATGTGTATCTA--AACCCAGCTCAGGTGTCTCCAGCTGT 2496
QY 243 ValHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGluSerLeu 262
DB 2497 GGATTTAAGGCAAAAGGCTTGACATCCACGATGATTAATCCCAATGCTTCCTTA 2556
QY 263 MetThrArgCysTrpSerLysAspProSerGlnArgProSerMetGluLysIleValLys 282
DB 2557 ATAGTGCGCTTGCTGGCCGATGAGCATGAAACGTCCTTTTCCAGCATTTATGAA 2616
QY 283 IleMetThrHisLeuMetArgTrpPhePro 292
DB 2617 ACCTTGAAACCAATGACTAAACAGCGCA 2646

RESULT 14
US-09-938-842A-1577
Sequence 1577, Application US/0993842A
Patent No. US20020160378A1
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Kreps, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
FILE REFERENCE: SCRIPT100-3
CURRENT APPLICATION NUMBER: US/09/938,842A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227,866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264,647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300,111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 1577
LENGTH: 2211
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-1577

Alignment Scores:
Pred. No.: 2,22e-36 Length: 2211
Score: 391.50 Matches: 98
Percent Similarity: 53.36% Conservative: 53
Best Local Similarity: 34.63% Mismatches: 99
Query Match: 24.24% Indels: 33
DB: 9 Gaps: 13

US-09-830-144-2_COPY_1_303 (1-303) x US-09-938-842A-1577 (1-2211)

QY 30 GluIleAspTrpLysGluIleGluValGluLysValGluValArgGlyAlaPheGlyVal 49
DB 1384 GAGATACGATGGAGAAAGATCTTACAACTTGGGAGAGAGTCGAAAGAGTTTATTTGTCGC 1443
QY 50 ValCysLysAlaLysTrpArgAlaLysAspValAlaIleLysGlnIleGluSerGluSer 69
DB 1444 GTTCATCGTCGAGTTTGGAAATGATCGCATCTGCTATTAA--GTTTACTTCGATGCT 1500
QY 70 GluArgLysAlaPheIleVal-----GluLeuArgGlnLeuSerArgVal 84
DB 1501 GATTACATGCGATGATCTTACGCGAGTCGAAAGAGATCAACATTATGAAGAACTG 1560
QY 85 AsnHisProAsnIleValLysLeuTrpGlyValAcysLeuAsnProValCys-----101
DB 1561 AGACATCCGAATGTGCTACTATTATGGAGCA-----GTATGTACAGAGAA 1608
QY 102 -----LeuValMetGluTrpAlaGluGlyGlySerLeuTrpAsnValIleGly 118
DB 1609 AAATCGCCATTAATCAAGAAATATATGCCAAGAGGAGTCTTCCAAATCTTCATAT 1668
QY 119 Ala--GluProLeuProTrpTrpThrAlaAlaHisAlaMetSerTrpCysLeuGlnCys 137
DB 1669 ACGAATCAGCATTCG-----GACAAAGAAAGCCGTTTAAATGCGCCTTGATGT 1719
QY 138 SerGlnGlyValAlaTrpLeuHisSerMetGlnProLysAlaLeuLeuHisArgAspLeu 157
DB 1720 GCTAGGGGAATGAAATTACTTACACCGCAGAAATCG--CCAAATTGATCATAGACTTG 1776
QY 158 LysProProAsnLeuLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGly 177
DB 1777 AAATCTTCATCTACTCTCGTGACCAAGAACTGGAATGTC--AAGTTGAGACTTTGGG 1833
QY 178 -----ThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySer 192
DB 1834 TTATCAAGTGAAGAAAGCAACCTTCTTGATGACT-----AAATCCGGAAAGGAAC 1887
QY 193 AlaAlaTrpMetAlaProGluValPheGluGlySerAsnTrpSerGlnLysCysAspVal 212
DB 1888 CCGCAGTGAATGGCTCTGAAAGTTCTCAGAAAGTGAACCTTCGATGATGAAGTGTG 1947
QY 213 PheSerTrpGlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIle 232
DB 1948 TTGAGCTTTGGAGTCATCTTATGAGAGCTAATGACTACGTTATACCATGACCGCTTG 2007
QY 233 GlyGlyProAlaPheArgIleMetTrpAlaVal--HisAsnGlyThrArgProProLeu 251
DB 2008 AAC-----TCTATTGAGTTGTTGGAGTTGTTGTTTCATGATGACGATTAAGACTTA 2061
QY 252 IleLysAsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspPro 271
DB 2062 CCTGAAGATTAATCCCGGATCGCATCAATATACAGAGATTGTGGCAACTGATCCA 2121
QY 272 SerGlnArgProSerMetGluLysIleValLysIleMetThrHisLeuMetArgTrpPhe 291
DB 2122 GCAAAACGACGTCGTCGAGGAATTAATGACATGATATGAGCCTGTTCCGCAAA--2178
QY 292 ProGlyAla 294
DB 2179 CCAAGGTCA 2187

RESULT 15
US-09-938-842A-882
Sequence 882, Application US/0993842A
Patent No. US20020160378A1
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Kreps, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
FILE REFERENCE: SCRIPT100-3
CURRENT APPLICATION NUMBER: US/09/938,842A

; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 882
; LENGTH: 1428
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-882

Alignment Scores:
Pred. No.: 3,89e-36 Length: 1428
Score: 387.00 Matches: 96
Percent Similarity: 52.20% Conservative: 58
Best Local Similarity: 32.54% Mismatches: 109
Query Match: 23.96% Indels: 32
DB: 9 Gaps: 10

US-09-830-144-2_COPY_1_303 (1-303) x US-09-938-842A-882 (1-1428)

```
QY 10 SerSerSerSerSerAlaGlyGluMetIleGluAlaProSerGlnValLeuAsnPheGlu 29
DB 427 TCTAAGTAAGTTCGCAGAGTGTGGAAGATGC-----462
QY 30 GluIleAspTyrLysGluIleGluValGluGluValValGlyArgGlyAlaPheGlyVal 49
DB 463 CTCATTGATGTCTTAAGTTGCTTATGGGATAGGTTTGTCTCACGGGAAATATAGCCAG 522
QY 50 ValCysLysAlaLysTrpArgAlaLysAspValAlaIleLysGlnIleGluSerGluSer 69
DB 523 ATTTATCATGGTGAATATGAAGCAAGCTGTGTCTCTGAAGATTATCACAGCGCTGAG 582
QY 70 Glu-----ArgLysAlaPheIleValGluLeuArg 79
DB 583 GATAGTGACGACATATCTTGGGAGCTCGTTAGAAAAAGAGTTTATCGTGGAGCCACT 642
QY 80 GlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyValaCysLeuAsnPro 99
DB 643 CTTTATCTCGACTAAGCCATCCAAATGTCGTTAAGTTTGTGGAGTGAATACTCGAAAC 702
QY 100 ValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnValLeuHisGlyAla 119
DB 703 ---TGTCATCATCAGAGATGTACCTCGAGGTCTTTAAGATCATATCTGCACCAAGCTC 759
QY 120 Glu-----ProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeuGlnCys 137
DB 760 GAGCAGAAATCCCTTCCT-----TTGGAACAGCTAATCGATTTTGGTCTGGATATT 810
QY 138 SerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeu 157
DB 811 GCTAAAGGAATGGAATATATTCACCTCA-----AGAGAGATAGTTTCATCAGGATCTG 861
QY 158 LysProProAsnLeuLeuValAlaGlyThrValLeuLysIleCysAspPheGly 177
DB 862 AAGCCAGAAAAC---GTGTTGATCGACAATGACTTTTCATCTGAAGATTCTGCATTTGGC 918
QY 178 ThrAlaCysAspIleGln-----ThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 195
DB 919 ATAGCGTGCAGGAGGAGTACTGTGATGTTTGGGGGATAACATAGGAACTTATAGGTGG 978
QY 196 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 215
DB 979 ATGGCACCTGAAGTTTAAACCGGATACCACATCGGAGGATGCGATGTTTATAGTTT 1038
QY 216 GlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 235
DB 1039 GGACTCTTTTATGGGAAATGGGTAGCTGGAGCACTTCCATATGAGGAGATGAATTT--- 1095
QY 236 AlaPheArgIleMetTrpAlaVal---HisAsnGlyThrArgProLeuIleLysAsn 254
```

```
DB 1096 GCTGAACAAATTGCCTACGCAGTTATATACAAGAAAATTAGCCAGTTATACCGACGGAT 1155
QY 255 LeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArg 274
DB 1156 TGTCCAGCGGCCCATGAAGAGCTGATCGAGCGATGTTGGTCATCGCAACACACAAGAGA 1215
QY 275 ProSerMetGluGluIleValLysIleMetThrHisLeuMetArg 289
DB 1216 CCGGAATTCTGGCAGATTGTCAAAGTGTGGAACATTTCAAGAAG 1260
```

Search completed: December 10, 2002, 03:46:27
Job time : 114 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 10, 2002, 03:48:55 ; Search time 37 Seconds
(without alignments)
821.112 Million cell updates/sec

Title: US-09-830-144-2_COPY_76_303

Perfect score: 1252
Sequence: 1 VELRLSRVNHPIVKLYCA.....MTHLMRYPPGADEPLQYPCQ 228

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A_Geneseq_101002.*
- 1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT.*
 - 2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT.*
 - 3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT.*
 - 4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT.*
 - 5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT.*
 - 6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT.*
 - 7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT.*
 - 8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT.*
 - 9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT.*
 - 10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT.*
 - 11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT.*
 - 12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT.*
 - 13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT.*
 - 14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT.*
 - 15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT.*
 - 16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT.*
 - 17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT.*
 - 18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT.*
 - 19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT.*
 - 20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT.*
 - 21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT.*
 - 22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT.*
 - 23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1252	100.0	567	20 AAY28998	Human TGF-beta act
2	1252	100.0	579	18 AAW27092	Mouse transforming
3	1252	100.0	579	18 AAW27093	Human transforming
4	1252	100.0	579	20 AAY28996	Human TGF-beta act
5	1252	100.0	579	20 AAY09542	Human TAK1 protein
6	1252	100.0	579	21 AAY91000	Human TAK-1 protei
7	1252	100.0	579	23 ABB85033	Pain regulated pro
8	1252	100.0	590	20 AAY09547	Human TAK1-GxHis p
9	1252	100.0	606	20 AAY28997	Human TGF-beta act
10	700	55.9	678	22 ABB58061	Drosophila melanog

11	414	33.1	252	22	ABB60985	Drosophila melanog
12	374	29.9	367	21	AAG32053	Arabidopsis thalia
13	374	29.9	369	21	AAG32052	Arabidopsis thalia
14	374	29.9	407	21	AAG32051	Arabidopsis thalia
15	373	29.8	369	21	AAG22172	Arabidopsis thalia
16	373	29.8	374	21	AAG22171	Arabidopsis thalia
17	373	29.8	412	21	AAG22170	Arabidopsis thalia
18	371	29.6	349	22	AAG75571	Human colon cancer
19	371	29.6	455	21	AA18657	A human regulator
20	371	29.6	455	21	ABY83378	Human survival reg
21	371	29.6	455	21	AA184321	A human cardiovasc
22	371	29.6	473	22	AAW25322	Human protein sequ
23	371	29.6	800	22	AAW71957	Human TGF-beta rec
24	371	29.6	800	22	AAW71957	Novel protein kina
25	363.5	29.0	1020	22	ABY58999	Drosophila melanog
26	355.5	28.4	1021	23	ABP61000	Novel human protei
27	355	28.4	719	22	AA85513	Human protein kina
28	355	28.4	1036	23	ABB80923	Novel human protei
29	353.5	28.2	859	16	AAW82886	Human leucine-zipp
30	353.5	28.2	859	18	AAW31127	Human leucine-zipp
31	352.5	28.2	888	23	ABY57049	Mouse ischaemic co
32	351.5	28.1	977	22	ABY71694	Drosophila melanog
33	350	28.0	589	21	AA45384	Arabidopsis thalia
34	350	28.0	732	21	AA45383	Arabidopsis thalia
35	350	28.0	760	21	AA45382	Arabidopsis thalia
36	349.5	27.9	1046	22	AAE11775	Human kinase (PKIN
37	349.5	27.9	1097	23	AAE21717	Human PKIN-12 prot
38	348	27.8	319	21	AAQ25601	Arabidopsis thalia
39	348	27.8	341	21	AAQ25600	Arabidopsis thalia
40	348	27.8	391	21	AAQ25599	Arabidopsis thalia
41	347	27.7	338	21	AAQ28422	Arabidopsis thalia
42	347	27.7	338	21	AAQ50301	Arabidopsis thalia
43	347	27.7	346	21	AAQ28421	Arabidopsis thalia
44	347	27.7	346	21	AAQ50300	Arabidopsis thalia
45	345.5	27.6	847	23	AAE22763	Human mitogen acti

ALIGNMENTS

RESULT 1

AA28998

ID AAY28998 standard; Protein; 567 AA.

XX AAY28998;

AC AAY28998;

DT 29-OCT-1999 (first entry)

XX Human TGF-beta activated kinase (TAK) 1c amino acid sequence.

DE Nuclear factor kappa B; NF-kB, inhibitor; TGF-beta activated kinase 1;

XX TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;

KW intracitable disease; atrophic dermatitis; psoriasis; viral infection;

KW endotoxin shock; septicemia; human; hTAK1c.

OS Homo sapiens.

XX WO9940202-A1.

PN 12-AUG-1999.

XX 02-FEB-1999; 99WO-JP00422.

XX 30-OCT-1998; 98JP-0309316.

PR 06-FEB-1998; 98JP-0026003.

XX (TANA) TANABE SEIVAKU CO.

XX Hasegawa K, Kageyama N, Sakurai H, Sugita T;

PI WPI; 1999-494298/41.

XX N-PSDB; AAX99698.

PT Nuclear factor kappa B activation inhibitors, useful as preventives
PT For, e.g. autoimmune diseases
PS Examples: Page 43-46; 49pp; Japanese.
XX

CC The invention provides a method for identifying or screening a nuclear
CC factor kappa B (NF-kB) activation inhibitor by examining the effect of a
CC test substance on modulating the function(s) of TGF-beta activated kinase
CC 1 (TAK1). The NFkB activation inhibitors targeting on TAK1 can be used to
CC treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis),
CC intractable diseases with inflammation (such as atrophic dermatitis and
CC psoriasis), viral infection, endotoxin shock, septicemia and others. The
CC present sequence represents the amino acid sequence of human TAK1c
CC (hTAK1c) protein.
CC
XX

SQ Sequence 567 AA;
Query Match 100.0%; Score 1252; DB 20; Length 567;
Best Local Similarity 100.0%; Pred. No.2.9e-127;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VELRQLSRVNHPNIVKLYGACLNPCVCLWMEYAGSLYVNLHGAEPLPYTTAAHMSWCL 60
DB 76 VELRQLSRVNHPNIVKLYGACLNPCVCLWMEYAGSLYVNLHGAEPLPYTTAAHMSWCL 135
QY 61 QCSQGVAYLHSMQPKALIHRLKPNLLVAGGTVLKICDFTGACDIQTHMTNNGSAAM 120
DB 136 QCSQGVAYLHSMQPKALIHRLKPNLLVAGGTVLKICDFTGACDIQTHMTNNGSAAM 195
QY 121 MAPEVFGSNYSKCDVFSNGIILMEVITRRKPFDEIGGPAFIMAVHNGTRPPLIKNL 180
DB 196 MAPEVFGSNYSKCDVFSNGIILMEVITRRKPFDEIGGPAFIMAVHNGTRPPLIKNL 255
QY 181 PKPIESLMTRCWSKDPSPRSMEIIVKIMTHLMRYFGADEPLQYPCQ 228
DB 256 PKPIESLMTRCWSKDPSPRSMEIIVKIMTHLMRYFGADEPLQYPCQ 303

RESULT 2
AAW27092
ID AAW27092 standard; Protein; 579 AA.
XX
AC AAW27092;
XX
DT 19-NOV-1997 (first entry)
XX
DE Mouse transforming growth factor-beta activated kinase TAK-1.
XX
KW TGF-beta; signal transduction; TGF-beta activated kinase;
KW MAPK kinase activator; AMK-1; bone morphogenetic protein; BMP;
KW protein kinase.
XX
OS Mus musculus.
XX
PN JP09163990-A.
XX
PD 24-JUN-1997.
XX
PF 27-SEP-1996; 96JP-0256747.
XX
PR 24-JUL-1996; 96US-0685625.
PR 29-SEP-1995; 95JP-0253549.
XX
PA (CHUS) CHUGAI PHARM CO LTD.
PA (UENO/) UENO N.
XX
DR WPI; 1997-380171/35.
DR N-PSDB; AAT85094.
XX
PT DNA encoding transforming growth factor-beta-activated kinase, TAK-1
PT - useful for studying the TGF-beta signal transduction system
XX
PS Claim 14; Page 10-12; 20pp; Japanese.

XX
CC The present sequence represents mouse transforming growth factor-beta
CC (TGF-beta) activated kinase, TAK-1. The DNA is used to produce the
CC TAK-1 protein which is involved in the TGF-beta family signal
CC transduction system. TAK-1, also known as activator of MAPK Kinase
CC (AMK-1), is an enzyme which is activated by TGF-beta and bone
CC morphogenetic protein (BMP) and activates MAPK kinase by
CC phosphorylation.
CC
XX

SQ Sequence 579 AA;
Query Match 100.0%; Score 1252; DB 18; Length 579;
Best Local Similarity 100.0%; Pred. No.2.9e-127;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VELRQLSRVNHPNIVKLYGACLNPCVCLWMEYAGSLYVNLHGAEPLPYTTAAHMSWCL 60
DB 76 VELRQLSRVNHPNIVKLYGACLNPCVCLWMEYAGSLYVNLHGAEPLPYTTAAHMSWCL 135
QY 61 QCSQGVAYLHSMQPKALIHRLKPNLLVAGGTVLKICDFTGACDIQTHMTNNGSAAM 120
DB 136 QCSQGVAYLHSMQPKALIHRLKPNLLVAGGTVLKICDFTGACDIQTHMTNNGSAAM 195
QY 121 MAPEVFGSNYSKCDVFSNGIILMEVITRRKPFDEIGGPAFIMAVHNGTRPPLIKNL 180
DB 196 MAPEVFGSNYSKCDVFSNGIILMEVITRRKPFDEIGGPAFIMAVHNGTRPPLIKNL 255
QY 181 PKPIESLMTRCWSKDPSPRSMEIIVKIMTHLMRYFGADEPLQYPCQ 228
DB 256 PKPIESLMTRCWSKDPSPRSMEIIVKIMTHLMRYFGADEPLQYPCQ 303

RESULT 3
AAW27093
ID AAW27093 standard; Protein; 579 AA.
XX
AC AAW27093;
XX
DT 19-NOV-1997 (first entry)
XX
DE Human transforming growth factor-beta activated kinase TAK-1.
XX
KW TGF-beta; signal transduction; TGF-beta activated kinase;
KW MAPK kinase activator; AMK-1; bone morphogenetic protein; BMP;
KW protein kinase.
XX
OS Homo sapiens.
XX
PN JP09163990-A.
XX
PD 24-JUN-1997.
XX
PF 27-SEP-1996; 96JP-0256747.
XX
PR 24-JUL-1996; 96US-0685625.
PR 29-SEP-1995; 95JP-0253549.
XX
PA (CHUS) CHUGAI PHARM CO LTD.
PA (UENO/) UENO N.
XX
DR WPI; 1997-380171/35.
DR N-PSDB; AAT85095.
XX
PT DNA encoding transforming growth factor-beta-activated kinase, TAK-1
PT - useful for studying the TGF-beta signal transduction system
XX
PS Claim 15; Page 13-15; 20pp; Japanese.
XX
CC The present sequence represents human transforming growth factor-beta
CC (TGF-beta) activated kinase, TAK-1. The DNA is used to produce the
CC TAK-1 protein which is involved in the TGF-beta family signal
CC transduction system. TAK-1, also known as activator of MAPK Kinase
CC (AMK-1), is an enzyme which is activated by TGF-beta and bone

CC morphogenetic protein (BMP) and activates MAPK kinase by
XX phosphorylation.

SQ Sequence 579 AA;

Query Match 100.0%; Score 1252; DB 18; Length 579;
Best Local Similarity 100.0%; Pred. No. 2.9e-127;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VELRQLSRVNHNPNIKLYGACLNPCVCLVMEYAEAGGSLYNVLHGAELPPYYTAHAHMSWCL 60
DB 76 VELRQLSRVNHNPNIKLYGACLNPCVCLVMEYAEAGGSLYNVLHGAELPPYYTAHAHMSWCL 135
QY 61 QCSQGVAYLHSMOPKALIHRLDKPPNLLVAGGTVLKICDFGTACDIQTHMTNNKGSAAW 120
DB 136 QCSQGVAYLHSMOPKALIHRLDKPPNLLVAGGTVLKICDFGTACDIQTHMTNNKGSAAW 195
QY 121 MAPEVFEFGSNYSEKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 180
DB 196 MAPEVFEFGSNYSEKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 255
QY 181 PKPIESLMTRCWSKDPQSRPSMEIIVKIMTHLMRYFPGADEPLOYPCQ 228
DB 256 PKPIESLMTRCWSKDPQSRPSMEIIVKIMTHLMRYFPGADEPLOYPCQ 303

RESULT 4

AAAY28996
ID AAY28996 standard; Protein; 579 AA.

AC AAY28996;

XX 29-OCT-1999 (first entry)

XX Human TGF-beta activated kinase (TAK) 1a amino acid sequence.

XX Nuclear factor kappa B; NF-kB; inhibitor; TGF-beta activated kinase 1;
KW TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;
KW intractable disease; atrophic dermatitis; psoriasis; viral infection;
KW endotoxin shock; septicemia; human; hTAK1a.

XX Homo sapiens.

XX WO9940202-A1.

XX 12-AUG-1999.

XX 02-FEB-1999; 99WO-JP00422.

XX 30-OCT-1998; 98JP-0309316.

PR 06-FEB-1998; 98JP-0026003.

XX (TANA) TANABE SEIYAKU CO.

XX Hasegawa K, Kageyama N, Sakurai H, Sugita T;

XX WPI; 1999-494298/41.

DR N-PSDB; AAX99696.

XX Nuclear factor kappa B activation inhibitors, useful as preventives
PT for, e.g. autoimmune diseases

XX Examples; Page 35-39; 49pp; Japanese.

XX The invention provides a method for identifying or screening a nuclear
CC factor kappa B (NF-kB) activation inhibitor by examining the effect of a
CC test substance on modulating the function(s) of TGF-beta activated kinase
CC 1 (TAK1). The NFkB activation inhibitors targeting on TAK1 can be used to
CC treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis),
CC intractable diseases with inflammation (such as atrophic dermatitis and
CC psoriasis), viral infection, endotoxin shock, septicemia and others. The
CC present sequence represents the amino acid sequence of human TAK1a
CC (hTAK1a) protein.

XX SQ Sequence 579 AA;

Query Match 100.0%; Score 1252; DB 20; Length 579;
Best Local Similarity 100.0%; Pred. No. 2.9e-127;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VELRQLSRVNHNPNIKLYGACLNPCVCLVMEYAEAGGSLYNVLHGAELPPYYTAHAHMSWCL 60
DB 76 VELRQLSRVNHNPNIKLYGACLNPCVCLVMEYAEAGGSLYNVLHGAELPPYYTAHAHMSWCL 135
QY 61 QCSQGVAYLHSMOPKALIHRLDKPPNLLVAGGTVLKICDFGTACDIQTHMTNNKGSAAW 120
DB 136 QCSQGVAYLHSMOPKALIHRLDKPPNLLVAGGTVLKICDFGTACDIQTHMTNNKGSAAW 195
QY 121 MAPEVFEFGSNYSEKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 180
DB 196 MAPEVFEFGSNYSEKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 255
QY 181 PKPIESLMTRCWSKDPQSRPSMEIIVKIMTHLMRYFPGADEPLOYPCQ 228
DB 256 PKPIESLMTRCWSKDPQSRPSMEIIVKIMTHLMRYFPGADEPLOYPCQ 303

RESULT 5

AAAY09542
ID AAY09542 standard; Protein; 579 AA.

AC AAY09542;

XX 21-JUL-1999 (first entry)

XX Human TAK1 protein.

XX Human; TAK1; TAK1; screening; inhibition; TGF-beta;
KW transforming growth factor beta.

XX Homo sapiens.

XX Key Location/Qualifiers
FT CDS 183..1922
FT /*tag= a

XX WO9921010-A1.

XX 29-APR-1999.

XX 22-OCT-1998; 98WO-JP04796.

XX 22-OCT-1997; 97JP-0290188.

XX (CHUS) CHUGAI SEIYAKU KK.

XX Ohtomo T, Ono K, Tsuchiya M;

XX WPI; 1999-312645/26.

DR N-PSDB; AAX56279.

XX Screening for TGF- beta inhibitory substances, which are useful as
PT drugs for treatment of diseases relating to its disorder

XX Claim 4; Page 155-157; 195pp; Japanese.

XX A method has been developed for screening for substances which inhibit
CC the binding of TAK1 polypeptide to TAK1 polypeptide. The method
CC comprises: (a) contacting the polypeptide in the presence of a sample;
CC and (b) detecting the amount of bound polypeptide, in which the sample
CC can be pre-mixed with TAK1 or TAK1 polypeptide first. The transforming
CC growth factor (TGF)-beta inhibitory substances can be used in drugs for
CC indications e.g. as TGF-beta signal transduction inhibitors or
CC activators, or extracellular matrix protein production enhancement
CC inhibitors or activators, or cell proliferation prevention inhibitors or
CC activators, or monocyte migration inhibitors or activators, or

CC physiological activity induction inhibitors or activators, or
CC immunosuppression inhibitors or activators, or amyloid beta protein
CC precipitation inhibitors or activators, and such substances can also be
CC inhibitors of the TAK1 polypeptide function, particularly kinase
CC activity. The present sequence represents human TAK1.

XX Sequence 579 AA;

Query Match 100.0%; Score 1252; DB 20; Length 579;
Best Local Similarity 100.0%; Pred. No. 2,9e-127;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VELRLSRVNHPIVLYGACINPVCLVMEYABGSLYVNLHGAEPPLYTAHAMSWCL 60
DB 76 VELRLSRVNHPIVLYGACINPVCLVMEYABGSLYVNLHGAEPPLYTAHAMSWCL 135
QY 61 QCSQGVAYLHSMQPKALLHRDLKPPNLLVAGTVLKICDFGTACDIQTHMTNNGSAAW 120
DB 136 QCSQGVAYLHSMQPKALLHRDLKPPNLLVAGTVLKICDFGTACDIQTHMTNNGSAAW 195
QY 121 MAPEVFECSNYSEKCDVFSWGIILMEVITRRKPFDEIGGPAFRIMAVANGTRPPLIKNL 180
DB 196 MAPEVFECSNYSEKCDVFSWGIILMEVITRRKPFDEIGGPAFRIMAVANGTRPPLIKNL 255
QY 181 PKPIESLMTRCWSKDPSPQSPMEIYKIMTHLMRYFPGADEPLQYPCQ 228
DB 256 PKPIESLMTRCWSKDPSPQSPMEIYKIMTHLMRYFPGADEPLQYPCQ 303

RESULT 6
AA91000
ID AA91000 standard; Protein; 579 AA.

XX AA91000;

DT 04-SEP-2000 (first entry)

DE Human TAK-1 protein sequence SEQ ID NO:2.

XX Human; TAK-1; TAB-1; mitogen activated protein kinase; MAPK;
XX screening; signal transduction; inhibition; inflammatory cytokine;
XX IL-1; interleukin 1; TNF; tumour necrosis factor; inflammation;
XX antiinflammatory; suppression.

OS Homo sapiens.

PN WO200023610-A1.

PD 27-APR-2000.

PF 21-OCT-1999; 99WO-JP05817.

PR 21-OCT-1998; 98JP-0299962.

PA (CHUS) CHUGAI SEIYAKU KK.

PI Tsuchiya M, Ohtomo T, Sugamata Y, Matsumoto K;

DR N-PSDB; AAA39105.

PT Method for screening inhibitors of TAK1 signal transduction for

PT suppression of inflammatory cytokine production and use as

PT antiinflammatory agents -

PS Example 1; Page 80-84; 100pp; Japanese.

XX The present invention describes a method for screening compounds for
CC inhibition of inflammatory cytokine signal transduction by contacting
CC the sample with TAK1 and its receptor TAB1 and selecting for inhibition
CC of TAK1/TAB1 binding. Also described is a method for screening compounds
CC for inhibition of inflammatory cytokine signal transduction in which the
CC inhibition of TAK1 phosphorylation is selected for; and drug

CC compositions for the treatment of inflammatory disorders containing as
CC active component an inflammatory cytokine signal transduction inhibitor.
CC TAK1 is an essential component of the signalling process which results
CC in release of inflammatory cytokines such as interleukin-1 (IL-1),
CC IL-10, tumour necrosis factor (TNF) and IL-6. The methods can be used
CC for the selection of effective antiinflammatory agents. The present
CC sequence represents human TAK-1, which is used in the exemplification of
CC the present invention.

XX Sequence 579 AA;

Query Match 100.0%; Score 1252; DB 21; Length 579;
Best Local Similarity 100.0%; Pred. No. 2,9e-127;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VELRLSRVNHPIVLYGACINPVCLVMEYABGSLYVNLHGAEPPLYTAHAMSWCL 60
DB 76 VELRLSRVNHPIVLYGACINPVCLVMEYABGSLYVNLHGAEPPLYTAHAMSWCL 135
QY 61 QCSQGVAYLHSMQPKALLHRDLKPPNLLVAGTVLKICDFGTACDIQTHMTNNGSAAW 120
DB 136 QCSQGVAYLHSMQPKALLHRDLKPPNLLVAGTVLKICDFGTACDIQTHMTNNGSAAW 195
QY 121 MAPEVFECSNYSEKCDVFSWGIILMEVITRRKPFDEIGGPAFRIMAVANGTRPPLIKNL 180
DB 196 MAPEVFECSNYSEKCDVFSWGIILMEVITRRKPFDEIGGPAFRIMAVANGTRPPLIKNL 255
QY 181 PKPIESLMTRCWSKDPSPQSPMEIYKIMTHLMRYFPGADEPLQYPCQ 228
DB 256 PKPIESLMTRCWSKDPSPQSPMEIYKIMTHLMRYFPGADEPLQYPCQ 303

RESULT 7

ABB85033
ID ABB85033 standard; Protein; 579 AA.

XX ABB85033;

DT 16-MAY-2002 (first entry)

DE Pain regulated protein sequence 28.

XX Pain; analgesic; gene therapy; neurological disorder;
XX neurodegenerative disease.

OS Homo sapiens.

PN WO200212338-A2.

PD 14-FEB-2002.

PF 03-AUG-2001; 2001WO-EP09011.

PR 03-AUG-2000; 2000DE-1037759.

PA (CHEF) GRUENENTHAL GMBH.

PI Gillen C, Wetzel S, Wenzel S, weine E, Schaefer MK;

DR N-PSDB; ABL88437.

PT Identifying pain-regulating compounds, useful for treating chronic pain

PT and for diagnosis, by measuring binding of compounds to specific

PT peptides and proteins -

PS Claim 1; Fig 44; 213pp; German.

XX The invention relates to identifying pain-regulating substances (A)
CC comprising (i) incubating a test substance with a cell (or preparation
CC from it) that has synthesised a peptide or protein (B) and (ii) measuring
CC either binding of the test substance to (B) or some functional parameter
CC that is altered by this binding. The method is useful for identifying

CC pain-regulating substances (A) with analgesic activity. (A) along with
 CC nucleic acid (AB188411-ABL88441) that encode proteins (B,
 CC ABB85006-ABB85037) that interact with (A); (B): vectors containing the
 CC nucleic acid; antibodies against (B); cells that express (B) and agents
 CC that bind to (B), are all useful for treating pain, particularly chronic
 CC pain, including use in gene therapy. The same materials can also be used
 CC for diagnosis, e.g. of neurological and neurodegenerative diseases. The
 CC present sequence is that of a polypeptide of the invention.

XX Sequence 579 AA;

SQ Query Match 100.0%; Score 1252; DB 23; Length 579;
 Best Local Similarity 100.0%; Pred. No. 2.9e-127;
 Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VELRLSRVNHNPVIVKLYGACLPVCLVMEYAEAGGSLYNVLHGAEPLPYTAAHMSWCL 60
 DB 76 VELRLSRVNHNPVIVKLYGACLPVCLVMEYAEAGGSLYNVLHGAEPLPYTAAHMSWCL 135
 QY 61 QCSQGVAYLHSMOPKALIHRLDKPPNLLVAGGTVLKICDFGTACDIOTHTMTNKGSAW 120
 DB 136 QCSQGVAYLHSMOPKALIHRLDKPPNLLVAGGTVLKICDFGTACDIOTHTMTNKGSAW 195
 QY 121 MAPEVFEAGSYSEKCDVFSWGIIWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 180
 DB 196 MAPEVFEAGSYSEKCDVFSWGIIWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 255
 QY 181 PKPIESLWTRCWSKQSPRSMEIEIVKIMTHLMRYFPFGADEPLOYPCQ 228
 DB 256 PKPIESLWTRCWSKQSPRSMEIEIVKIMTHLMRYFPFGADEPLOYPCQ 303

RESULT 8
 AAY09547
 ID AAY09547 standard; Protein; 590 AA.

XX AC AAY09547;
 DT 21-JUL-1999 (first entry)
 XX DE Human TAK1-6xHis protein.
 XX Human; TAB1; screening; inhibition; TGF-beta;
 KW transforming growth factor beta.

XX OS Homo sapiens.
 OS Synthetic.

XX PN WO9921010-A1.

XX PD 29-APR-1999.

XX PF 22-OCT-1998; 98WO-JF04796.

XX PR 22-OCT-1997; 97JP-0290188.

XX PA (CHUS) CHUGAI SEIYAKU KK.

XX PI Ohtomo T, Ono K, Teuchiya M;

XX DR WPI; 1999-312645/26.

XX DR N-PSDB; AAX56285.

XX Screening for TGF- beta inhibitory substances, which are useful as
 PT drugs for treatment of diseases relating to its disorder

XX Example 1; Page 171-174; 195pp; Japanese.

XX A method has been developed for screening for substances which inhibit
 CC the binding of TAK1 polypeptide to TAB1 polypeptide. The method
 CC comprises: (a) contacting the polypeptide in the presence of a sample;
 CC and (b) detecting the amount of bound polypeptide, in which the sample
 CC can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming

CC growth factor (TGF)-beta inhibitory substances can be used in drugs for
 CC indications e.g. as TGF-beta signal transduction inhibitors or
 CC activators, or extracellular matrix protein production enhancement
 CC inhibitors or activators, or cell proliferation prevention inhibitors or
 CC activators, or monocyte migration inhibitors or activators, or
 CC physiological activity induction inhibitors or activators, or
 CC immunosuppression inhibitors or activators, or amyloid beta protein
 CC precipitation inhibitors or activators, and such substances can also be
 CC inhibitors of the TAK1 polypeptide function, particularly kinase
 CC activity. The present sequence represents TAK1-6xHis from an example of
 CC the present invention.

XX SQ Sequence 590 AA;

Query Match 100.0%; Score 1252; DB 20; Length 590;
 Best Local Similarity 100.0%; Pred. No. 3e-127;
 Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VELRLSRVNHNPVIVKLYGACLPVCLVMEYAEAGGSLYNVLHGAEPLPYTAAHMSWCL 60
 DB 76 VELRLSRVNHNPVIVKLYGACLPVCLVMEYAEAGGSLYNVLHGAEPLPYTAAHMSWCL 135
 QY 61 QCSQGVAYLHSMOPKALIHRLDKPPNLLVAGGTVLKICDFGTACDIOTHTMTNKGSAW 120
 DB 136 QCSQGVAYLHSMOPKALIHRLDKPPNLLVAGGTVLKICDFGTACDIOTHTMTNKGSAW 195
 QY 121 MAPEVFEAGSYSEKCDVFSWGIIWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 180
 DB 196 MAPEVFEAGSYSEKCDVFSWGIIWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 255
 QY 181 PKPIESLWTRCWSKQSPRSMEIEIVKIMTHLMRYFPFGADEPLOYPCQ 228
 DB 256 PKPIESLWTRCWSKQSPRSMEIEIVKIMTHLMRYFPFGADEPLOYPCQ 303

RESULT 9
 AAY28997
 ID AAY28997 standard; Protein; 606 AA.

XX AC AAY28997;

XX DT 29-OCT-1999 (first entry)

XX DE Human TGF-beta activated kinase (TAK) 1b amino acid sequence.

XX KW Nuclear factor kappa B; NF-kB; inhibitor; TGF-beta activated kinase 1;
 KW TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;
 KW intractable disease; atrophic dermatitis; psoriasis; viral infection;
 KW endotoxin shock; septicemia; human; hTAK1b.

XX OS Homo sapiens.

XX PN WO9940202-A1.

XX PD 12-AUG-1999.

XX PF 02-FEB-1999; 99WO-JF00422.

XX PR 30-OCT-1998; 98JP-0309316.

XX PR 06-FEB-1998; 98JP-0026003.

XX PA (TANA) TANABE SEIYAKU CO.

XX PI Hasegawa K, Kageyama N, Sakurai H, Sugita T;

XX DR WPI; 1999-494298/41.

XX DR N-PSDB; AAX99697.

XX Nuclear factor kappa B activation inhibitors, useful as preventives
 PT for, e.g. autoimmune diseases

XX Examples; Page 39-43; 49pp; Japanese.

CC The invention provides a method for identifying or screening a nuclear
CC factor kappa B (NF-kB) activation inhibitor by examining the effect of a
CC test substance on modulating the function(s) of TGF-beta activated kinase
CC 1 (TAK1). The NFkB activation inhibitors targeting on TAK1 can be used to
CC treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis),
CC intractable diseases with inflammation (such as atrophic dermatitis and
CC psoriasis), viral infection, endotoxin shock, septicemia and others. The
CC present sequence represents the amino acid sequence of human TAK1b
CC (hTAK1b) protein.
CC
SQ Sequence 606 AA;
XX
XX
Query Match 100.0%; Score 1252; DB 20; Length 606;
Best Local Similarity 100.0%; Pred. No. 3,1e-127;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VELRLSRVNHNPVIVKLYGACLNPCVLMVEYAGGSLYVNLHGAEPLPYTAAHMSMCL 60
DB 76 VELRLSRVNHNPVIVKLYGACLNPCVLMVEYAGGSLYVNLHGAEPLPYTAAHMSMCL 135
QY 61 QCSQGVAYIHSNQPKALIRDLKPPNLLVAGTVLKITCDFTACDIQTHMTNKGSAW 120
DB 136 QCSQGVAYIHSNQPKALIRDLKPPNLLVAGTVLKITCDFTACDIQTHMTNKGSAW 195
QY 121 MAPEVEGSGNSYSEKCDVFSMGIIIMEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 180
DB 196 MAPEVEGSGNSYSEKCDVFSMGIIIMEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 255
QY 181 PKRIESLMTRCWSKDPSPQSPSMEIIVKIMTHLMRYFPGADEPLQYPCQ 228
DB 256 PKRIESLMTRCWSKDPSPQSPSMEIIVKIMTHLMRYFPGADEPLQYPCQ 303
RESULT 10
ABB58061
ID ABB58061 standard; Protein; 678 AA.
XX
AC ABB58061;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 975.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
DR N-PSDB; ABL02164.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 975; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL01840-ABL16175), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ffp.wipo.int/pub/published_pct_sequences.
CC
SQ Sequence 678 AA;
XX
XX
Query Match 55.9%; Score 700; DB 22; Length 678;
Best Local Similarity 57.5%; Pred. No. 4.9e-67;
Matches 130; Conservative 35; Mismatches 57; Indels 4; Gaps 3;
QY 2 ELRLSRVNHNPVIVKLYG--ACLNPCVLMVEYAGGSLYVNLHGAEPLPYTAAHMSMC 59
DB 60 EVKQLSRVNHNPVIVKLYG--ACLNPCVLMVEYAGGSLYVNLHGAEPLPYTAAHMSMC 118
QY 60 LQCSQGVAYIHSNQPKALIRDLKPPNLLVAGTVLKITCDFTACDIQTHMTNKGSAW 119
DB 119 RQCAEGLAYIHAMTPKPLIRHDVPLNLLITNKGRLKITCDFTVADKSTMTNKGSAW 178
QY 120 MAPEVEGSGNSYSEKCDVFSMGIIIMEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 179
DB 179 MAPEVEGSGNSYSEKCDVFSMGIIIMEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 237
QY 180 LKPKIESLMTRCWSKDPSPQSPSMEIIVKIMTHLMRYFPGADEPLQY 225
DB 238 CPKRIEDLMTACWKVYEDRPSMGYIVGVHVEIKDYTGADKALEY 283
RESULT 11
ABB60985
ID ABB60985 standard; Protein; 252 AA.
XX
AC ABB60985;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 9747.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
DR N-PSDB; ABL05088.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 9747; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of

[illegible]

PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136031.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140595.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144086.
PR 16-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 06-OCT-1999; 99US-0157753.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.

```
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

Query Match 29.9%; Score 374; DB 21; Length 369;

Best Local Similarity 39.4%; Pred. No. 6, 5e-32; Mismatches 70; Indels 22; Gaps 9;

Matches 86; Conservative 40; Mismatches 70; Indels 22; Gaps 9;

```
QY 2 ELRQLSRVNHPIVVKLYGACLNPNV--CLVMEYAEGSGLYNVL----HGAEPLPYTAHA 55
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 137 EYSMNLNLKHPNIVRFICGRKPMWCVITEYAKGGSVHQFLTRQRNRAVPLKL-----A 191
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 56 MSNCLCCSGCVATLHSMOKALIHRLKRPNNLLVAGTVLKCIPRGTA-CDIOTH-MTN 113
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 192 VKQALDVARGMAVYHG--RNFLHRLKSDNLLISADKSI-KIADGVARIIVQTEGMT 247
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 114 NKGSAAWMAPEVEGSGNYSEKCDVFSMGIIILWEVITRRKPFDEIGG--PAFRIMAVHNG 171
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 248 ECTTYMMAPEMQHRAYNQKDVYSFGIVLWELITGLLPQNMTAVQAFAFV--VVRG 304
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 172 TRPPLIKNLPKPIESLMTRCWSDPSQPSMEETVKIM 209
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 305 VRPTVPNDCLPVLSIDIMTRCWDANPEVRPCFEVVKLL 342
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

RESULT 14

AAG32051

ID AAG32051 standard; Protein; 407 AA.

XX AAG32051;

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 38595.

KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

XX Arabidopsis thaliana.

OS EPI033405-A2.

PN 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

```
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 11-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138647.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139763.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140981.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144335.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
```


PR 19-JUL-1999;	99US-0144335.	PR 14-OCT-1999;	99US-0159330.
PR 20-JUL-1999;	99US-0144352.	PR 14-OCT-1999;	99US-0159331.
PR 20-JUL-1999;	99US-0144632.	PR 14-OCT-1999;	99US-0159637.
PR 20-JUL-1999;	99US-0144884.	PR 14-OCT-1999;	99US-0159638.
PR 21-JUL-1999;	99US-0144814.	PR 18-OCT-1999;	99US-0159584.
PR 21-JUL-1999;	99US-0145086.	PR 21-OCT-1999;	99US-0160741.
PR 21-JUL-1999;	99US-0145088.	PR 21-OCT-1999;	99US-0160767.
PR 22-JUL-1999;	99US-0145085.	PR 21-OCT-1999;	99US-0160768.
PR 22-JUL-1999;	99US-0145087.	PR 21-OCT-1999;	99US-0160770.
PR 22-JUL-1999;	99US-0145089.	PR 21-OCT-1999;	99US-0160814.
PR 22-JUL-1999;	99US-0145192.	PR 21-OCT-1999;	99US-0160815.
PR 23-JUL-1999;	99US-0145114.	PR 22-OCT-1999;	99US-0160980.
PR 23-JUL-1999;	99US-0145218.	PR 22-OCT-1999;	99US-0160981.
PR 23-JUL-1999;	99US-0145224.	PR 22-OCT-1999;	99US-0160989.
PR 26-JUL-1999;	99US-0145278.	PR 25-OCT-1999;	99US-0161404.
PR 27-JUL-1999;	99US-0145913.	PR 25-OCT-1999;	99US-0161405.
PR 27-JUL-1999;	99US-0145918.	PR 25-OCT-1999;	99US-0161406.
PR 27-JUL-1999;	99US-0145919.	PR 26-OCT-1999;	99US-0161359.
PR 28-JUL-1999;	99US-0145951.	PR 26-OCT-1999;	99US-0161360.
PR 02-AUG-1999;	99US-0146386.	PR 26-OCT-1999;	99US-0161361.
PR 02-AUG-1999;	99US-0146388.	PR 28-OCT-1999;	99US-0161920.
PR 02-AUG-1999;	99US-0146389.	PR 28-OCT-1999;	99US-0161992.
PR 03-AUG-1999;	99US-0147038.	PR 28-OCT-1999;	99US-0161993.
PR 04-AUG-1999;	99US-0147204.	PR 29-OCT-1999;	99US-0162142.
PR 05-AUG-1999;	99US-0147302.		
PR 05-AUG-1999;	99US-0147260.		
PR 06-AUG-1999;	99US-0147303.		
PR 06-AUG-1999;	99US-0147416.		
PR 09-AUG-1999;	99US-0147493.		
PR 09-AUG-1999;	99US-0147935.		
PR 10-AUG-1999;	99US-0148171.		
PR 11-AUG-1999;	99US-0148319.		
PR 12-AUG-1999;	99US-0148341.		
PR 13-AUG-1999;	99US-0148565.		
PR 13-AUG-1999;	99US-0148684.		
PR 16-AUG-1999;	99US-0149368.		
PR 17-AUG-1999;	99US-0149175.		
PR 18-AUG-1999;	99US-0149426.		
PR 20-AUG-1999;	99US-0149722.		
PR 20-AUG-1999;	99US-0149723.		
PR 20-AUG-1999;	99US-0149929.		
PR 23-AUG-1999;	99US-0149902.		
PR 23-AUG-1999;	99US-0149930.		
PR 25-AUG-1999;	99US-0150566.		
PR 26-AUG-1999;	99US-0150884.		
PR 27-AUG-1999;	99US-0151065.		
PR 27-AUG-1999;	99US-0151066.		
PR 27-AUG-1999;	99US-0151080.		
PR 30-AUG-1999;	99US-0151303.		
PR 31-AUG-1999;	99US-0151438.		
PR 01-SEP-1999;	99US-0151930.		
PR 07-SEP-1999;	99US-0152363.		
PR 10-SEP-1999;	99US-0153070.		
PR 13-SEP-1999;	99US-0153758.		
PR 15-SEP-1999;	99US-0154018.		
PR 16-SEP-1999;	99US-0154039.		
PR 20-SEP-1999;	99US-0154779.		
PR 22-SEP-1999;	99US-0155139.		
PR 23-SEP-1999;	99US-0155486.		
PR 24-SEP-1999;	99US-0155659.		
PR 28-SEP-1999;	99US-0156458.		
PR 29-SEP-1999;	99US-0156596.		
PR 04-OCT-1999;	99US-0157117.		
PR 05-OCT-1999;	99US-0157753.		
PR 06-OCT-1999;	99US-0157865.		
PR 07-OCT-1999;	99US-0158029.		
PR 08-OCT-1999;	99US-0158232.		
PR 12-OCT-1999;	99US-0158369.		
PR 13-OCT-1999;	99US-0159293.		
PR 13-OCT-1999;	99US-0159294.		
PR 13-OCT-1999;	99US-0159295.		
PR 14-OCT-1999;	99US-0159329.		
PR 19-JUL-1999;	99US-0144335.	Query Match	29.9%; Score 374; DB 21; Length 407;
PR 20-JUL-1999;	99US-0144352.	Best Local Similarity	39.4%; Pred. No. 7.4e-32;
PR 20-JUL-1999;	99US-0144632.	Matches	86; Conservative 40; Mismatches 70; Indels 22; Gaps 9;
PR 20-JUL-1999;	99US-0144884.		
PR 21-JUL-1999;	99US-0144814.		
PR 21-JUL-1999;	99US-0145086.		
PR 21-JUL-1999;	99US-0145088.		
PR 22-JUL-1999;	99US-0145085.		
PR 22-JUL-1999;	99US-0145087.		
PR 22-JUL-1999;	99US-0145089.		
PR 22-JUL-1999;	99US-0145192.		
PR 23-JUL-1999;	99US-0145114.		
PR 23-JUL-1999;	99US-0145218.		
PR 23-JUL-1999;	99US-0145224.		
PR 26-JUL-1999;	99US-0145278.		
PR 27-JUL-1999;	99US-0145913.		
PR 27-JUL-1999;	99US-0145918.		
PR 27-JUL-1999;	99US-0145919.		
PR 28-JUL-1999;	99US-0145951.		
PR 02-AUG-1999;	99US-0146386.		
PR 02-AUG-1999;	99US-0146388.		
PR 02-AUG-1999;	99US-0146389.		
PR 03-AUG-1999;	99US-0147038.		
PR 04-AUG-1999;	99US-0147204.		
PR 05-AUG-1999;	99US-0147302.		
PR 05-AUG-1999;	99US-0147260.		
PR 06-AUG-1999;	99US-0147303.		
PR 06-AUG-1999;	99US-0147416.		
PR 09-AUG-1999;	99US-0147493.		
PR 09-AUG-1999;	99US-0147935.		
PR 10-AUG-1999;	99US-0148171.		
PR 11-AUG-1999;	99US-0148319.		
PR 12-AUG-1999;	99US-0148341.		
PR 13-AUG-1999;	99US-0148565.		
PR 13-AUG-1999;	99US-0148684.		
PR 16-AUG-1999;	99US-0149368.		
PR 17-AUG-1999;	99US-0149175.		
PR 18-AUG-1999;	99US-0149426.		
PR 20-AUG-1999;	99US-0149722.		
PR 20-AUG-1999;	99US-0149723.		
PR 20-AUG-1999;	99US-0149929.		
PR 23-AUG-1999;	99US-0149902.		
PR 23-AUG-1999;	99US-0149930.		
PR 25-AUG-1999;	99US-0150566.		
PR 26-AUG-1999;	99US-0150884.		
PR 27-AUG-1999;	99US-0151065.		
PR 27-AUG-1999;	99US-0151066.		
PR 27-AUG-1999;	99US-0151080.		
PR 30-AUG-1999;	99US-0151303.		
PR 31-AUG-1999;	99US-0151438.		
PR 01-SEP-1999;	99US-0151930.		
PR 07-SEP-1999;	99US-0152363.		
PR 10-SEP-1999;	99US-0153070.		
PR 13-SEP-1999;	99US-0153758.		
PR 15-SEP-1999;	99US-0154018.		
PR 16-SEP-1999;	99US-0154039.		
PR 20-SEP-1999;	99US-0154779.		
PR 22-SEP-1999;	99US-0155139.		
PR 23-SEP-1999;	99US-0155486.		
PR 24-SEP-1999;	99US-0155659.		
PR 28-SEP-1999;	99US-0156458.		
PR 29-SEP-1999;	99US-0156596.		
PR 04-OCT-1999;	99US-0157117.		
PR 05-OCT-1999;	99US-0157753.		
PR 06-OCT-1999;	99US-0157865.		
PR 07-OCT-1999;	99US-0158029.		
PR 08-OCT-1999;	99US-0158232.		
PR 12-OCT-1999;	99US-0158369.		
PR 13-OCT-1999;	99US-0159293.		
PR 13-OCT-1999;	99US-0159294.		
PR 13-OCT-1999;	99US-0159295.		
PR 14-OCT-1999;	99US-0159329.		
Qy 2 ELRQLSRVHPNIVKLYGACLNIV-CLWMEYAEGLSLNVL-----HGAEPLEPYTAHA 55			
Db 175 EVSMLANKHPNIVFTIGACRKPVMVCIIVTEYAKGGSVRQFLTRQNRVAPLKL-----A 229			
Qy 56 MSWCLQCSQGVAYLHSMQPKALIHRLKPPNLLLVAGGTVLKICDFGTA-CDIQTH-MTN 113			
Db 230 VKQALDVARGMAIVHG---RNFHRLKSNLLISADKSI-KIADFGVARIEVQTEGMP 285			
Qy 114 NKGSAAMWAPFEGSNYSKCDVFSWGIILWEVITRRKPFDEIGG--PAPRIMWAVHNG 171			
Db 286 ETGYRWMAPEMIQHRAYNQKVDVYSGIVLWELITGLLPQNMTAVQAAFAV---VNRG 342			
Qy 172 TRPLIKNLKPFTESLMTKRCWSDPSPORPSMEIIVKIM 209			
Db 343 VRPTVPNDCLPVLSDIMTRCWDANPEVRPCFVGVKLL 380			
RESULT 15			
AAAG22172			
ID	AAAG22172 standard; Protein; 369 AA.		
XX			
AC	AAAG22172;		
XX			
DT	17-OCT-2000 (first entry)		
XX			
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 24999.		
XX			
KW	Protein identification; signal transduction pathway; metabolic pathway;		
KW	hybridisation assay; genetic mapping; gene expression control; promoter;		
KW	termination sequence.		
XX			
OS	Arabidopsis thaliana.		
XX			
PN	EP1033405-A2.		
XX			
PD	06-SEP-2000.		
XX			
PF	25-FEB-2000; 2000EP-0301439.		
XX			
PR	25-FEB-1999; 99US-0121825.		
PR	05-MAR-1999; 99US-0123180.		
PR	09-MAR-1999; 99US-0123548.		
PR	23-MAR-1999; 99US-0125788.		
PR	25-MAR-1999; 99US-0126264.		

PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135253.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140685.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142330.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144086.
PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149422.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149920.
PR 26-AUG-1999; 99US-0150566.
PR 27-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 15-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154039.
PR 22-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0156559.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.

```
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 29.8%; Score 373; DB 21; Length 369;
Best Local Similarity 39.9%; Pred. No. 8.3e-32;
Matches 87; Conservative 39; Mismatches 70; Indels 22; Gaps 9;

QY 2 ELRQLSRVNHNPVIVKLYGACLNVP--CLWMEYAEAGSLYNVL-----HGAELPYTTAHA 55
Db 137 EVSMLAFLKHPNIVRFIGACIKPMVWCIVTEYAKGGSVRQFLTKRQRAVPLKL-----A 191

QY 56 MSWCQCQSGVAYLHSMQPKALIHRLKPPNLLVAGGTVLKICDFGTA-CDIOTH-MTN 113
Db 192 VMQALDVARGMAYVHE--RNFIHRDLKSDNLLISADRSI-KIADFGVARIETEGMTP 247

QY 114 NKGSAAMWAPVEFGSNGSEKDFVSWGIILWEVITRRKPFDEIGG--PAPRIMWAVHNG 171
Db 248 ETGTYRWAPEMIQHPPTQKVDDVYSFGIVLWELITGLLPFQNMATAVQAFAV---VNRG 304

QY 172 TRPPLIKNLKPKIESLMTFCWSKDPQRSMBEIVKIM 209
Db 305 VRPTVPADCLPVLGEIMTRCWDADPEVRPCFAEIVNLL 342
```

Search completed: December 10, 2002, 03:57:13
Job time : 38 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 10, 2002, 03:52:45 ; Search time 20 Seconds
(without alignments)
1095.932 Million cell updates/sec

Title: US-09-830-144-2_COPY_76_303
Perfect score: 1252
Sequence: 1 VELRQLSRVNHNPVVKLYGA.....MTHLMRYFPGADEPLQYPCQ 228

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 73: *
1: Pir1: *
2: Pir2: *
3: Pir3: *
4: Pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1252	100.0	567	JC5957	transforming growt
2	1252	100.0	579	JC5955	transforming growt
3	1252	100.0	606	JC5956	transforming growt
4	392.5	31.3	954	S68178	mixed-lineage prot
5	374	29.9	407	G84635	probable protein k
6	373	29.8	412	T10671	protein kinase hom
7	363.5	29.0	394	JU0229	mixed-lineage prot
8	362	28.9	1338	T18287	protein-tyrosine k
9	356	28.4	462	S29851	protein kinase 6 (
10	355.5	28.4	370	T46150	protein kinase ATN
11	353.5	28.2	668	JC2363	protein kinase (BC
12	353.5	28.2	888	A53118	serine/threonine p
13	351.5	28.1	888	JC5399	dual leucine zippe
14	350	28.0	546	D84555	probable protein k
15	348	27.8	391	T48115	protein kinase ATN
16	347	27.7	475	T12955	probable protein k
17	346	27.6	328	T16747	hypothetical prote
18	346	27.6	1030	F96763	hypothetical prote
19	345.5	27.6	553	T04683	mixed-lineage prot
20	345.5	27.6	847	A53800	protein kinase hom
21	344	27.5	390	T01451	protein-tyrosine k
22	341	27.2	410	B35670	hypothetical prote
23	340	27.2	387	T22511	hypothetical prote
24	336	26.8	356	S61766	protein kinase ATN
25	329.5	26.3	357	C84856	probable protein k
26	329.5	26.3	345	T05675	hypothetical prote
27	328.5	26.2	377	T46149	protein kinase ATN
28	328.5	26.2	736	T05137	protein kinase hom
29	324	25.9	982	T06576	probable protein k

ALIGNMENTS

RESULT 1

JC5957
transforming growth factor-beta activated kinase (EC 2.7.-.-) lc - human
C:Species: Homo sapiens (man)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 17-Mar-2000
C:Accession: JC5957
R:Sakurai, H.; Shigemori, N.; Hasegawa, K.; Sugita, T.
Biochem. Biophys. Res. Commun. 243, 545-549, 1998
A:Title: TGF-beta-activated kinase 1 stimulates NF-kappaB activation by an NF-kappaB-ind
A:Reference number: JC5955; MUID:98153801; PMID:9480845
A:Accession: JC5957
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-567 <SA>
A:Cross-references: DBJ:AB009358
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C:Keywords: phosphotransferase

Query Match	100.0%	Score 1252;	DB 2;	Length 567;
Best Local Similarity	100.0%;	Pred. No. 1.5e-72;		
Matches 228;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	VELRQLSRVNHNPVVKLYGACLNPCVLVMEYABGGSLYNVLHGAEPLPYTTAAHMSWCL	60	
Db	76	VELRQLSRVNHNPVVKLYGACLNPCVLVMEYABGGSLYNVLHGAEPLPYTTAAHMSWCL	135	
Qy	61	QCSQGVAYLHSMQPKALIHRLKPPNLLVAGGTVLKICDFGTACDIQTHMTNKGSAAW	120	
Db	136	QCSQGVAYLHSMQPKALIHRLKPPNLLVAGGTVLKICDFGTACDIQTHMTNKGSAAW	195	
Qy	121	MAPEVFGSNYSKCDVFSWGIIWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL	180	
Db	196	MAPEVFGSNYSKCDVFSWGIIWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL	255	
Qy	181	PKPIESLMTFCWSDQSPQSRPSMBEIVKIMTHLMRYFPGADEPLQYPCQ	228	
Db	256	PKPIESLMTFCWSDQSPQSRPSMBEIVKIMTHLMRYFPGADEPLQYPCQ	303	

RESULT 2

JC5955
transforming growth factor-beta activated kinase (EC 2.7.-.-) 1a - human
C:Species: Homo sapiens (man)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: JC5955
R:Sakurai, H.; Shigemori, N.; Hasegawa, K.; Sugita, T.
Biochem. Biophys. Res. Commun. 243, 545-549, 1998
A:Title: TGF-beta-activated kinase 1 stimulates NF-kappaB activation by an NF-kappaB-ind
A:Reference number: JC5955; MUID:98153801; PMID:9480845
A:Accession: JC5955
A:Status: preliminary
A:Molecule type: DNA

30	323	25.8	356	2	T48206	protein kinase ATN
31	323	25.8	821	2	T48400	serine/threonine-p
32	321.5	25.7	364	2	G71410	probable protein k
33	320	25.6	829	2	T07406	probable protein k
34	319.5	25.5	438	2	C86273	protein kinases ho
35	319.5	25.5	445	2	T31581	hypothetical prote
36	315.5	25.2	1064	1	S57450	protein-tyrosine k
37	314	25.1	406	2	T52626	probable mitogen-a
38	313	25.0	1015	2	T00726	protein f33b2.2 (i
39	308.5	24.6	848	2	B87950	hypothetical prote
40	308.5	24.6	855	2	T20082	protein-tyrosine k
41	307.5	24.6	1115	1	S29926	protein-tyrosine k
42	304.5	24.3	402	2	B34735	protein-tyrosine k
43	304.5	24.3	435	2	JN0290	protein-tyrosine k
44	304.5	24.3	477	2	JN0291	transforming prote
45	304.5	24.3	801	4	TVHURE	

A:Residues: 1-579 <SAK>
A:Cross-references: DDBJ:AB009356; NID:g2924623; PIDN:BA025025.1; PID:g2924624
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C:Keywords: phosphotransferase

Query Match 100.0%; Score 1252; DB 2; Length 579;
Best Local Similarity 100.0%; Pred. No. 1.5e-72;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VELRLSRVNHNPVIVKLYGACINPVCVMEYAEAGSLYVNLHGAEPPLPYTAAHMSMCL 60
DB 76 VELRLSRVNHNPVIVKLYGACINPVCVMEYAEAGSLYVNLHGAEPPLPYTAAHMSMCL 135

QY 61 QCSQGVAVLHSMQPKALHRLDKPNNLLVAGTVLKI CDGTICDIQTHTNNKGSAAW 120
DB 136 QCSQGVAVLHSMQPKALHRLDKPNNLLVAGTVLKI CDGTICDIQTHTNNKGSAAW 195

QY 121 MAPEVEGNSYSEKCDVFSWGIIIMEVITRRKPFDEIGGPARIMAAVHNGTRPPLIKNL 180
DB 196 MAPEVEGNSYSEKCDVFSWGIIIMEVITRRKPFDEIGGPARIMAAVHNGTRPPLIKNL 255

QY 181 PKPIESLMTKRCWSKDPSPQPSMEIIVKIMTHLMRYFPGADEPLQYPCQ 228
DB 256 PKPIESLMTKRCWSKDPSPQPSMEIIVKIMTHLMRYFPGADEPLQYPCQ 303

RESULT 3
JC5956
transforming growth factor-beta activated kinase (EC 2.7.-.-) 1b - human
C:Species: Homo sapiens (man)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: JC5956
R:Sakurai, H.; Shigemori, N.; Hasegawa, K.; Sugita, T.
Biochem. Biophys. Res. Commun. 243:545-549, 1998
A:Title: TGF-beta-activated kinase 1 stimulates NF-kappaB activation by an NF-kappaB-ind
A:Reference number: JC5955; MUID:98153801; PMID:9480845
A:Accession: JC5956
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-606 <SAK>
A:Cross-references: DDBJ:AB009357; NID:g2924625; PIDN:BA025026.1; PID:g2924626
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C:Keywords: phosphotransferase

Query Match 100.0%; Score 1252; DB 2; Length 606;
Best Local Similarity 100.0%; Pred. No. 1.6e-72;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VELRLSRVNHNPVIVKLYGACINPVCVMEYAEAGSLYVNLHGAEPPLPYTAAHMSMCL 60
DB 76 VELRLSRVNHNPVIVKLYGACINPVCVMEYAEAGSLYVNLHGAEPPLPYTAAHMSMCL 135

QY 61 QCSQGVAVLHSMQPKALHRLDKPNNLLVAGTVLKI CDGTICDIQTHTNNKGSAAW 120
DB 136 QCSQGVAVLHSMQPKALHRLDKPNNLLVAGTVLKI CDGTICDIQTHTNNKGSAAW 195

QY 121 MAPEVEGNSYSEKCDVFSWGIIIMEVITRRKPFDEIGGPARIMAAVHNGTRPPLIKNL 180
DB 196 MAPEVEGNSYSEKCDVFSWGIIIMEVITRRKPFDEIGGPARIMAAVHNGTRPPLIKNL 255

QY 181 PKPIESLMTKRCWSKDPSPQPSMEIIVKIMTHLMRYFPGADEPLQYPCQ 228
DB 256 PKPIESLMTKRCWSKDPSPQPSMEIIVKIMTHLMRYFPGADEPLQYPCQ 303

RESULT 4
S68178
mixed-lineage protein kinase 2 (EC 2.7.1.-.-) - human
C:Species: Homo sapiens (man)
C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 11-Jun-1999
C:Accession: S68178; I38044; S32468
R:Dorow, D.S.; Devereux, L.; Th, G.F.; Price, G.; Nicholl, J.K.; Sutherland, G.R.; Simps
Eur. J. Biochem. 234: 492-500, 1995

A:Title: Complete nucleotide sequence, expression, and chromosomal localisation of huma
A:Reference number: S68178; MUID:96128179; PMID:8536694
A:Accession: S68178
A:Molecule type: mRNA
A:Residues: 1-954 <DOR>
A:Cross-references: EMBL:X0846; NID:g9711419; PIDN:CAA62351.1; PID:g9711420
R:Katon, M.; Hirai, M.; Sugimura, T.; Terada, M.
Oncogene 10, 1447-1451, 1995
A:Title: Cloning and characterization of MST, a novel (putative) serine/threonine kinas
A:Reference number: I38044; MUID:95249256; PMID:7731697
A:Accession: I38044
A:Status: translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-461, 'A', 'V', '465-470', 'S', '472-806', 'R', '808-817', 'A', '819-954 <RES>
A:Cross-references: EMBL:Z48615; NID:g758592; PIDN:CAA8531.1; PID:g758593
R:Dorow, D.S.; Devereux, L.; Dietzsch, E.; de Kretser, T.
Eur. J. Biochem. 213, 701-710, 1993
A:Title: Identification of a new family of human epithelial protein kinases containing t
A:Reference number: S32467; MUID:93238756; PMID:8477742
A:Accession: S32468
A:Molecule type: mRNA
A:Residues: 244-464, 'AOAGRRPQPALWL' <DD2>
C:Genetics:
A:Gene: GDB:MLK2; GDB:MST
A:Cross-references: GDB:36264; GDB:624810; OMIM:600137
A:Map position: 19q13.1-19q13.2
C:Superfamily: mixed-lineage protein kinase 2; protein kinase homology; SH3 homology
C:Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific protein k
F:23-76/Domain: SH3 homology <SH3>
F:96-364/Domain: protein kinase homology <KIN>
F:104-112/Region: protein kinase ATP-binding motif
F:284-405/Region: leucine zipper motif
F:419-440/Region: leucine zipper motif
F:449-463/Region: basic
F:125,145,222,224/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 31.3%; Score 392.5; DB 1; Length 954;
Best Local Similarity 40.6%; Pred. No. 7.8e-18;
Matches 88; Conservative 35; Mismatches 77; Indels 17; Gaps 6;

QY 2 ELRLSRVNHNPVIVKLYGACINP--VCLMEYAEAGSLYVNLHGAEPPLPYTAAHMSMCL 59
DB 145 ELRLSRVNHNPVIVKLYGACINP--VCLMEYAEAGSLYVNLHGAEPPLPYTAAHMSMCL 100

QY 60 QCSQGVAVLHSMQPKALHRLDKPNNLLV-----AGTVLKI CDGTACD-IQTHM 111
DB 201 VQYARGMNYLHNDAPVPIIHRDLKSIINILIEAENHNLADTVLKITDGLAREHKKTK 260

QY 112 TNNKGSAAWMAPEVEGNSYSEKCDVFSWGIIIMEVITRRKPFDEIGGPARIMAAV-HN 170
DB 261 MSAGTYAMMAPEVEIRLSFSKSDVWSFGVLWELTGEVYFRI--DALAVAGVAMN 318

QY 171 GTRPPLIKNLPKPIESLMTKRCWSKDPSPQPSMEIIVK 207
DB 319 KLTLPISCTCEPFPARLLECCMDPDRHPRDGSILK 355

RESULT 5
G84635
Probable protein kinase (imported) - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 17-May-2002
C:Accession: G84635
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Niernan, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: AB04420; MUID:20083487; PMID:10617197
A:Accession: G84635
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-407 <STO>

A;Cross-references: GB:AE002093; NID:g4337195; PIDN:RAD18109.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g24360
A;Map position: 2
C;Superfamily: kinase-related transforming protein; protein kinase homology

Query Match 29.9%; Score 374; DB 2; Length 407;
Best Local Similarity 39.4%; Pred. No. 5.6e-17;
Matches 86; Conservative 40; Mismatches 70; Indels 22; Gaps 9;

QY 2 ELRQLSRVNHENIVKLYGACLNVP--CLVMEYAEAGGSLYNVL-----HGAEPLPYTTAAHA 55
Db 175 EVSMLANLKHENIVRFIGACRKPWMCIVTEYAKGSVRQFLTRQRNAVPLKL-----A 229

QY 56 MSWCLOCSQGVAYLHSMQPKALIHRLDKPPNLLVAGTVLKI CDGTA-CDIQTH-WTN 113
Db 230 VKQALDVARGMAYVHG-----RNFIHRLKSDNLLISADKSI-KIADFGVARIIEVQTEGMP 285

QY 114 NKGSAAWMAPEVFGSGNSYSEKCDVFSWGIIWEVITRRKPFDEIGG--PAFRIMWAVHG 171
Db 286 ETGTYRWMAPEMIQHRAYNQKVDVYSGVILWELITGLLPQNMTAVQAFAV---VNRG 342

QY 172 TRPPLIKNLKPKIESLMTRCWSKDPSPRSMEEIVKIM 209
Db 343 VRPTVPNDCLPVLSDIMTRCWDANPEVRPCFVGVVKLL 380

RESULT 6
T10671
protein kinase homolog F6E21.90 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 19-May-2000
C;Accession: T10671
R;Bevan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancroft
submitted to the Protein Sequence Database, June 1999
A;Reference number: Z16533
A;Accession: T10671
A;Molecule type: DNA
A;Residues: 1-412 <BEV>
A;Cross-references: EMBL:AL049914; GSPDB:GN00062; ATSP:F6E21.90
A;Experimental source: cultivar Columbia; BAC clone F6E21
C;Genetics:
A;Gene: ATSP:F6E21.90
A;Map position: 4
A;Introns: 300/2
C;Superfamily: kinase-related transforming protein; protein kinase homology
F:135-392/Domain: protein kinase homology <KIN>

Query Match 29.8%; Score 373; DB 2; Length 412;
Best Local Similarity 39.9%; Pred. No. 6.6e-17;
Matches 87; Conservative 39; Mismatches 70; Indels 22; Gaps 9;

QY 2 ELRQLSRVNHENIVKLYGACLNVP--CLVMEYAEAGGSLYNVL-----HGAEPLPYTTAAHA 55
Db 180 EVSMLAFLKHPNIVRFIGACIKPMWMCIVTEYAKGSVRQFLTKRQNAVPLKL-----A 234

QY 56 MSWCLOCSQGVAYLHSMQPKALIHRLDKPPNLLVAGTVLKI CDGTA-CDIQTH-WTN 113
Db 235 VKQALDVARGMAYVHE-----RNFIHRLKSDNLLISADRSI-KIADFGVARIIEVQTEGMP 290

QY 114 NKGSAAWMAPEVFGSGNSYSEKCDVFSWGIIWEVITRRKPFDEIGG--PAFRIMWAVHG 171
Db 291 ETGTYRWMAPEMIQHRPTQKVDVYSGVILWELITGLLPQNMTAVQAFAV---VNRG 347

QY 172 TRPPLIKNLKPKIESLMTRCWSKDPSPRSMEEIVKIM 209
Db 348 VRPTVPNDCLPVLGEINTRCWDADPEVRPCFABIVNLL 385

RESULT 7
JU0229
mixed-lineage protein kinase 1 - human
C;Species: Homo sapiens (man)

C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 17-Mar-1999
C;Accession: S32467; JU0229
R;Dorow, D.S.; Devereux, L.; Dietzsch, E.; de Kretser, T.
Eur. J. Biochem. 213, 701-710, 1993
A;Title: Identification of a new family of human epithelial protein kinases containing tv
A;Reference number: S32467; MUID:93238756; PMID:8477742
A;Accession: S32467
A;Molecule type: mRNA
A;Residues: 1-394 <DO2>
C;Genetics:
A;Gene: GDB:MLK1
A;Cross-references: GDB:141921; OMIM:600136
A;Map position: 14q24.3-14q31
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homology
C;Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific protein K;
F:1-269/Domain: protein kinase homology <KIN>
F:1-268/Domain: catalytic <CAT>
F:9-17/Region: protein kinase ATP-binding motif
F:289-310/Region: leucine zipper motif
F:324-345/Region: leucine zipper motif
F:354-368/Region: basic

Query Match 29.0%; Score 363.5; DB 2; Length 394;
Best Local Similarity 37.7%; Pred. No. 2.5e-16;
Matches 83; Conservative 36; Mismatches 84; Indels 17; Gaps 6;

QY 2 ELRQLSRVNHENIVKLYGACLNVP--VCLVMEYAEAGGSLYNVLHGAEPLPYTTAAHMSWC 59
Db 50 EAKLFAMLKHPNIIALRGVCLKEPNLCVMEFARGGPLNVLGSKRIPPD1----LVNWA 105

QY 60 LQCSQGVAYLHSMQPKALIHRLDKPPNLLV-----AGGTVLKICDRGTACD-IQTHM 111
Db 106 VQIARGMNVLDHDAIVPIIHRDLKSSNIILOQKVENGLSNKILKITDFGLAREWHRTTK 165

QY 112 TNNKGSAAWMAPEVFGSGNSYSEKCDVFSWGIIWEVITRRKPFDEIGGPAFRIMWAV-HN 170
Db 166 MSAAGTYWMAPEVIRASFSGSDVMSVGVLLWELLTGEVFRGIDG--LRVAYGVAMN 223

QY 171 GTRPPLIKNLKPKIESLMTRCWSKDPSPRSMEEIVKIMT 210
Db 224 KLALPSTCTPEPFAKLMECDWNPDPHSRPSFTNILDQLT 263

RESULT 8
T18287
protein-tyrosine kinase (EC 2.7.1.112) - slime mold (Dictyostelium discoideum)
C;Species: Dictyostelium discoideum
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Feb-2000
C;Accession: T18287
R;Adler, K.
submitted to the EMBL Data Library, July 1996
A;Reference number: Z18856
A;Accession: T18287
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1338 <ADL>
A;Cross-references: EMBL:U64830; NID:g1468982; PID:g1468983; PIDN:AAB04999.1
A;Introns: 1181/3
C;Keywords: ATP; phosphotransferase; tyrosine-specific protein kinase

Query Match 28.9%; Score 362; DB 2; Length 1338;
Best Local Similarity 38.1%; Pred. No. 9e-16;
Matches 80; Conservative 41; Mismatches 73; Indels 16; Gaps 6;

QY 2 ELRQLSRVNHENIVKLYGACLNVPVCLVMEYAEAGGSLYNVLH---GAEPYPYTTAAHMS 57
Db 1100 EVSSLKSHHPNVVTFMGARIDPPCIFTEYLQGGSLLDVLIHQIKIKLNPLMMYKWIHDL- 1158

QY 58 WCLQCSQGVAYLHSMQPKALIHRLDKPPNLLVAGTVLKI CDGTA-CDIQTHM-NNKG 116
Db 1159 -----SLGMEHLHSIQ---MLHRDLTSKNILLDEFKNI-KIADFGLATTLSDDDMTLSGIT 1209

submitted to the Protein Sequence Database, April 2000
A:Reference number: Z24459

A:Accession: T48115
A:Molecule type: DNA

A:Residues: 1-391 <RIE>

A:Cross-references: EMBL:AL138648; NID:g7523397; PIDN:CA86427.1; PID:g7523408
A:Experimental source: cultivar Columbia; BAC clone F16M2

R:Richimura, K.; Mizoguchi, T.; Shinozaki, K.

Plant Sci. 130, 171-179, 1997

A:Title: AtMKKI, an Arabidopsis protein kinase related to mammal mixed-lineage kinases a

A:Reference number: Z24427

A:Accession: T51942

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-391 <ICH>

A:Cross-references: EMBL:AB006810; NID:g2351096; PIDN:BA422079.1; PID:g2351097

A:Experimental source: cultivar Columbia

C:Genetics:

A:Map position: 3

A:Introns: 109/1; 149/3; 220/2; 278/3; 323/3

A>Note: F16M2.110

C:Superfamily: Kinase-related transforming protein; protein kinase homology

C:Keywords: phosphotransferase

Query Match 27.8% Score 348; DB 2; Length 391;
Best Local Similarity 33.7% Pred. No. 2.4e-15;
Matches 82; Conservative 35; Mismatches 76; Indels 50; Gaps 8;

QY 2 ELRQLSRVNHPIVKLYGACL-----NPVCLVMEYAEAGSGLY 38

Db 135 EVAVMQKLDHPNVTKFIGASMGTSDLRIPPAGDTGGRNGAHPARACCVVEYVAGTLLK 194

QY 39 NVLHGAEPILPYTA-----AHAMSWCLQCSQGVAYLHSMQPKALIRDLKPPNLLLVAGGT 94

Db 195 KFL-----IKKYRAKLPIKDVITQLALDLARGLSYLHS--KAIVHRDVXSENMLLOPNKT 246

QY 95 VLKICDPGTA---CDIQHTMTNNGSAAMAPEVEFGSGNSYSEKCDVFSFGIILWEVITRR 151

Db 247 -LKINADFGVARVEAQNPDQMTGETGTGLGYMAPEVLEKPKYRKCDVYSFGVCLWEIYICCD 305

QY 152 KP-----FDEIGGPAFRIMMAVHNGTRPPLIKNLKPKPIESLMTRCWXPSPQRPSEMEIV 206

Db 306 MPYADCSFAEISHA-----VVHRNLRPEIPKCCPHAVANIMKRCWDPNDRPEMEVEV 359

QY 207 KIM 209

Db 360 KLL 362

Search completed: December 10, 2002, 03:58:57
Job time : 21 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 10, 2002, 03:50:45 ; Search time 16 Seconds
(without alignments)
591.038 Million cell updates/sec

Title: US-09-830-144-2_COPY_76_303
Perfect score: 1252
Sequence: 1 VELQLSRVNHPIVKLYGA.....MTHLMRYPPGADEPLQYPCQ 228

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 segs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1252	100.0	579	1 M3K7 MOUSE	Q62073 mus musculus
2	1252	100.0	606	1 M3K7_HUMAN	Q43318 homo sapien
3	435.5	34.8	393	1 M3K7 DROME	P83104 drosophila
4	392.5	31.3	954	1 M3KA HUMAN	Q02779 homo sapien
5	363.5	29.0	394	1 M3K9 HUMAN	P80192 homo sapien
6	355.5	28.4	888	1 M3K3 RAT	Q63796 rattus norv
7	353.5	28.2	859	1 M3KC HUMAN	Q12852 homo sapien
8	353.5	28.2	888	1 M3KC_MOUSE	Q60700 mus musculus
9	341	27.2	410	1 KYK2 DICDI	P18161 dictyosteli
10	323	25.8	821	1 CTR1 ARATH	Q05609 arabidopsis
11	307.5	24.6	1115	1 RET_MOUSE	P35546 mus musculus
12	304.5	24.3	1114	1 RET_HUMAN	P07949 homo sapien
13	301.5	24.1	630	1 TEC_MOUSE	P24604 mus musculus
14	301.5	24.1	631	1 TEC_HUMAN	P42680 homo sapien
15	299.5	23.9	625	1 ITK_MOUSE	Q03526 mus musculus
16	298.5	23.8	620	1 ITK_HUMAN	Q08881 homo sapien
17	294.5	23.5	2347	1 KROS_HUMAN	P08922 homo sapien
18	292.5	23.4	402	1 KROS_AVISU	P00529 avian sarco
19	292	23.3	507	1 KROS_CHICK	P08941 gallus gall
20	290	23.2	1584	1 KYKI DICDI	P18160 dictyosteli
21	288.5	23.0	628	1 TESK RAT	Q63572 rattus norv
22	288	23.0	1307	1 PHYI CERPU	P25948 ceratodon p
23	286	22.8	804	1 FPS DROME	P18106 drosophila
24	284.5	22.7	832	1 ANR3_HUMAN	P57078 homo sapien
25	283.5	22.6	527	1 TXK_MOUSE	P42882 mus musculus
26	280	22.4	822	1 FGRI RAT	Q04589 rattus norv
27	279	22.3	467	1 MATK RAT	P41243 rattus norv
28	278.5	22.2	590	1 BTKL DROME	P08630 drosophila
29	278	22.2	822	1 FGRI_HUMAN	P11362 homo sapien
30	278	22.2	822	1 FGRI_MOUSE	P16092 mus musculus
31	275	22.0	984	1 EPA3 RAT	Q08680 rattus norv
32	274.5	21.9	801	1 FGR3_MOUSE	Q61851 mus musculus
33	271.5	21.7	1062	1 CC7_SCHPO	P41892 schizosacch

34	271	21.6	604	1 KRAA_MOUSE	P04627 mus musculus
35	271	21.6	604	1 KRAA_RAT	P14056 rattus norv
36	270.5	21.6	503	1 HCK_RAT	P50545 rattus norv
37	270	21.6	806	1 FGR3_HUMAN	P22607 homo sapien
38	269.5	21.5	507	1 MATK_HUMAN	P42679 homo sapien
39	269.5	21.5	626	1 TESK_HUMAN	Q15569 homo sapien
40	269	21.5	533	1 FPS_AVISU	P00541 avian sarco
41	269	21.5	819	1 FGRI_CHICK	P21804 gallus gall
42	269	21.5	983	1 EPA3_HUMAN	P29320 homo sapien
43	268	21.4	830	1 PKN2_MYXXA	P54736 myxococci
44	268	21.4	1052	1 FAK1_HUMAN	Q05397 homo sapien
45	268	21.4	1052	1 FAK1_MOUSE	P34152 mus musculus

ALIGNMENTS

RESULT 1					
M3K7_MOUSE					
ID	M3K7_MOUSE	STANDARD;	PRT;	579 AA.	
AC	Q62073;				
DT	16-OCT-2001 (Rel. 40, Created)				
DT	16-OCT-2001 (Rel. 40, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Mitogen-activated protein kinase kinase kinase 7 (EC 2.7.1.-)				
DE	(transforming growth factor-beta-activated kinase 1) (TGF-beta-activated kinase 1).				
GN	MAP3K7 OR TAK1.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				
OX	NCBI_TaxID=10090;				
RP	[1]				
RN	SEQUENCE FROM N.A.				
RX	MEDLINE=96123277; PubMed=8533096;				
RA	Yamaguchi K., Shirakabe K., Shibuya H., Irie K., Ohishi I., Ueno N.,				
RA	Taniguchi T., Nishida E., Matsumoto K.;				
RT	"Identification of a member of the MAPKKK family as a potential				
RT	mediator of TGF-beta signal transduction.";				
RL	Science 270:2008-2011(1995).				
CC	-!- FUNCTION: CAN PHOSPHORYLATE AND ACTIVATE YET UNDEFINED MAPKKS.				
CC	MEDIATOR OF TGF-BETA SIGNAL TRANSDUCTION. STIMULATES NF-KAPPA B				
CC	ACTIVATION.				
CC	-!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.				
CC	MAP KINASE KINASE SUBFAMILY.				
CC	-!- SIMILARITY: IN THE C-TERMINAL SECTION; STRONG, TO C21ORF7.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
CC	EMBL; D76446; BAA1184.1; -				
DR	HSSP; P08631; 1AD5.				
DR	MGI; 1346877; Map3k7.				
DR	InterPro; IPR000719; Euk_pkinase.				
DR	InterPro; IPR004040; STY_pkinase.				
DR	InterPro; IPR002290; Ser_thr_pkinase.				
DR	Pfam; PF00069; pkinase; 1.				
DR	ProDom; PD000001; Euk_pkinase; 1.				
DR	SMART; SM00221; STYKc; 1.				
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.				
DR	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.				
DR	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.				
KW	Transferase; Serine/threonine-protein kinase; ATP-binding.				
FT	DOMAIN 8 16 POLY-SER.				
FT	DOMAIN 36 291 PROTEIN_KINASE.				
FT	NP_BIND 42 50 ATP (BY SIMILARITY).				
FT	BINDING 63 63 ATP (BY SIMILARITY).				
FT	ACT_SITE 156 156 BY SIMILARITY.				

```

SQ SEQUENCE 579 AA, 64227 MW, 97C8F6F3C8E283EE CRC64;
Query Match 100.0%; Score 1252; DB 1; Length 579;
Best Local Similarity 100.0%; Pred. No. 3.7e-111;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VELRQLSRVNHPIVLYGACLNPCVLNMEYAEGLVNVHLGAEPLPYTTAAHMSWCL 60
DB 76 VELRQLSRVNHPIVLYGACLNPCVLNMEYAEGLVNVHLGAEPLPYTTAAHMSWCL 135
QY 61 QCSQGVAYLHSMQPKALIHRLDKPPLLVAAGTVLKCIDFGTACDIQTHMTNNGSAAM 120
DB 136 QCSQGVAYLHSMQPKALIHRLDKPPLLVAAGTVLKCIDFGTACDIQTHMTNNGSAAM 195
QY 121 MAPEVEGNSYSKCDVFSWGIIIMWEVITRRKPFDEIGGAPARIMAVANGTRPPLIKNL 180
DB 196 MAPEVEGNSYSKCDVFSWGIIIMWEVITRRKPFDEIGGAPARIMAVANGTRPPLIKNL 255
QY 181 PKPIESLMTRCWSKDPSPQSPSMEIYKIMTHLMRFPGADEPLQYPCQ 228
DB 256 PKPIESLMTRCWSKDPSPQSPSMEIYKIMTHLMRFPGADEPLQYPCQ 303

RESULT 2
M3K7 HUMAN
ID M3K7 HUMAN STANDARD; PRT; 606 AA.
AC 04318; 04317; 04319;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Mitogen-activated protein kinase kinase 7 (EC 2.7.1.-)
DE (Transforming growth factor-beta-activated kinase 1) (TGF-beta-
DE activated kinase 1).
GN MAPK7 OR TAK1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC TISSUE=Lung;
RX MEDLINE=98153801; Pubmed=9480845;
RA Sakurai H., Shigemori N., Hasegawa K., Sugita T.;
RA "TGF-beta-activated kinase 1 stimulates NF-kappa B activation by an
RT NF-kappa B-inducing kinase-independent mechanism.";
RL Biochem. Biophys. Res. Commun. 243:545-549(1998).
CC -1- FUNCTION: CAN PHOSPHORYLATE AND ACTIVATE YET UNDEFINED MARKS.
CC MEDIATOR OF TGF-BETA SIGNAL TRANSDUCTION. STIMULATES NF-KAPPA B
CC ACTIVATION.
CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1A, 1B (SHOWN HERE) AND 1C; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE KINASE SUBFAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; STRONG, TO C21ORF7.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB009357; BAA25026.1; -
DR EMBL; AB009356; BAA25025.1; -
DR EMBL; AB009358; BAA25027.2; -
DR HSSP; P08631; 1AD5.
DR Genew; HGNC:6859; MAP3K7.
DR MIM; 602614; -
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR004040; STR_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
```

```

DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00221; STYKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
KM Transferrase, Serine/threonine-protein kinase; ATP-binding;
KM Alternative splicing.
FT DOMAIN 8 14 POLY-SER.
FT NP_BIND 36 291 PROTEIN KINASE.
FT NP_BIND 42 50 ATP (BY SIMILARITY).
FT BINDING 63 63 ATP (BY SIMILARITY).
FT ACT_SITE 156 156 BY SIMILARITY.
FT VARSPIC 404 430 MISSING (IN ISOFORM 1A).
FT VARSPIC 509 518 PLAPCNSKE -> ARTSCRTSPG (IN ISOFORM 1C).
FT VARSPIC 519 606 MISSING (IN ISOFORM 1C).
SQ SEQUENCE 606 AA; 67196 MW; 3DBF8147CD174013 CRC64;

Query Match 100.0%; Score 1252; DB 1; Length 606;
Best Local Similarity 100.0%; Pred. No. 3.9e-111;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VELRQLSRVNHPIVLYGACLNPCVLNMEYAEGLVNVHLGAEPLPYTTAAHMSWCL 60
DB 76 VELRQLSRVNHPIVLYGACLNPCVLNMEYAEGLVNVHLGAEPLPYTTAAHMSWCL 135
QY 61 QCSQGVAYLHSMQPKALIHRLDKPPLLVAAGTVLKCIDFGTACDIQTHMTNNGSAAM 120
DB 136 QCSQGVAYLHSMQPKALIHRLDKPPLLVAAGTVLKCIDFGTACDIQTHMTNNGSAAM 195
QY 121 MAPEVEGNSYSKCDVFSWGIIIMWEVITRRKPFDEIGGAPARIMAVANGTRPPLIKNL 180
DB 196 MAPEVEGNSYSKCDVFSWGIIIMWEVITRRKPFDEIGGAPARIMAVANGTRPPLIKNL 255
QY 181 PKPIESLMTRCWSKDPSPQSPSMEIYKIMTHLMRFPGADEPLQYPCQ 228
DB 256 PKPIESLMTRCWSKDPSPQSPSMEIYKIMTHLMRFPGADEPLQYPCQ 303

RESULT 3
M3K7 DROME
ID M3K7 DROME STANDARD; PRT; 393 AA.
AC P83104;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative mitogen-activated protein kinase kinase 7
DE (EC 2.7.1.-).
DE TAK1.
GN Drosophila melanogaster (Fruit fly).
OS Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; Pubmed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.R., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballou R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokov D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durkin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
```


DR Pfam; PF00018; SH3; 1.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00452; SH3DOMAIN.
 DR PRINTS; PR00109; TYRKINASE.
 DR Prodom; PD000001; Euk_pkinase.
 DR Prodom; PD000066; SH3; 1.
 DR SMART; SM00326; SH3; 1.
 DR SMART; SM00321; STYK; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS50002; SH3; 1.
 DR Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase;
 KM ATP-binding; SH3 domain.
 FT DOMAIN 2 5
 FT DOMAIN 16 81
 FT DOMAIN 98 360
 FT NP_BIND 104 112
 FT BINDING 125 125
 FT ACT_SITE 222 222
 FT DOMAIN 384 405
 FT DOMAIN 419 440
 FT DOMAIN 449 463
 FT CONFLICT 462 464
 FT CONFLICT 465 480
 FT CONFLICT 471 471
 FT CONFLICT 807 807
 FT CONFLICT 818 818
 SQ SEQUENCE 954 AA; 103623 MW; 538F4AA559B0ABA CRC64;

Query Match 31.3%; Score 392.5; DB 1; Length 954;
 Best Local Similarity 40.6%; Pred. No. 2.1e-29;
 Matches 88; Conservative 35; Mismatches 77; Indels 17; Gaps 6;

QY 2 ELRQLSRVNHPIVVKLYGACLP--VCLVMEYAEAGSLYNNVHGAEPLPYTAHAHMSWC 59
 DB 145 EARLFGLAHPNIIALRGACINPPLCLVMEYARAGALRYLAGRVPVPHV---LVNWA 200
 QY 60 LQCSQGVAVILHSMOPKALIHDLKPNLLV-----AGGTVLKICDPGTACD-IQTHM 111
 DB 201 VQVARGMYLHNDAPVPIIHRDLKSNILILIEAHNNLADTLKIDTDFGLAREWHRTTK 260
 QY 112 TNNKGSAAWMAPEVEFGSYSEKCVFSWGIIILWEVITRRKPFDEIGGPARIMAV-HN 170
 DB 261 MSAAGTYAMMAPEVIRLSLFKSSDVWSFGVLWELLLGEPVREI--DALAVGVAMN 318
 QY 171 GTRPPLIKNLKPRIEISLMTRCWSKDPSPQSPMEIIVK 207
 DB 319 KLTLPFPTCPEPFAKLEECWDPDPHGRPDGFSILK 355

RESULT 5
 M3K9_HUMAN STANDARD; PRT; 394 AA.
 ID M3K9_HUMAN
 AC P80192;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Mitogen-activated protein kinase kinase 9 (EC 2.7.1.-) (Mixed
 DE lineage kinase 1) (Fragment).
 GN MAP3K9 OR MEK1 OR PRKE1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OC NCBI_TaxID=9606;
 RN 11
 RP SEQUENCE FROM N.A.
 RC TISSUE=Colon epithelium;
 RX MEDLINE=93238756; PubMed=8477742;
 RA Dorow D.S., Devereux L., Dietzsch E., de Kretser T.,
 RT "Identification of a new family of human epithelial protein kinases
 containing two leucine/isoleucine-zipper domains.";

RL Eur. J. Biochem. 213:701-710(1993).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN EPITHELIAL TUMOR CELL LINES OF
 CC COLONIC, BREAST AND GASTROPHAGEAL ORIGIN.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC MAP KINASE KINASE SUBFAMILY.
 DR PIR; S32467; S32467.
 DR PIR; J00229; J00229.
 DR HSSP; P12931; 1EMK.
 DR Genew; HGNC:6861; MAP3K9.
 DR MIM; 600136;
 DR Interpro; IPR00719; Euk_pkinase.
 DR Interpro; IPR002290; Ser_thr_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR Prodom; PD000001; Euk_pkinase; 1.
 DR SMART; SM00219; TYRK; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 KM Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase;
 KM ATP-binding.
 FT NON_TER 1 1
 FT DOMAIN 3 271
 FT NP_BIND 9 17
 FT BINDING 30 30
 FT ACT_SITE 127 127
 FT DOMAIN 289 310
 FT DOMAIN 324 345
 FT DOMAIN 354 368
 SQ SEQUENCE 394 AA; 44975 MW; DBE40B7D31047FD8 CRC64;

Query Match 29.0%; Score 363.5; DB 1; Length 394;
 Best Local Similarity 37.7%; Pred. No. 4.3e-27;
 Matches 83; Conservative 36; Mismatches 84; Indels 17; Gaps 6;

QY 2 ELRQLSRVNHPIVVKLYGACLP--VCLVMEYAEAGSLYNNVHGAEPLPYTAHAHMSWC 59
 DB 50 EARLFGLAHPNIIALRGACINPPLCLVMEYARAGALRYLAGRVPVPHV---LVNWA 105
 QY 60 LQCSQGVAVILHSMOPKALIHDLKPNLLV-----AGGTVLKICDPGTACD-IQTHM 111
 DB 106 VQVARGMYLHNDAPVPIIHRDLKSNILILIEAHNNLADTLKIDTDFGLAREWHRTTK 165
 QY 112 TNNKGSAAWMAPEVEFGSYSEKCVFSWGIIILWEVITRRKPFDEIGGPARIMAV-HN 170
 DB 166 MSAAGTYAMMAPEVIRLSLFKSSDVWSFGVLWELLLGEPVREI--DALAVGVAMN 223
 QY 171 GTRPPLIKNLKPRIEISLMTRCWSKDPSPQSPMEIIVK 210
 DB 224 KLTLPFPTCPEPFAKLEECWDPDPHGRPDGFSILK 263

RESULT 6
 M3K9_RAT STANDARD; PRT; 888 AA.
 ID M3K9_RAT
 AC Q63756;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Mitogen-activated protein kinase kinase 12 (EC 2.7.1.37)
 DE (MAPK-upstream kinase) (MUK).
 GN MAP3K12 OR MUK.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxID=10116;
 RN 11
 RP SEQUENCE FROM N.A.
 RC MEDLINE=96226099; PubMed=8637721;
 RX Hirai S., Izawa M., Osada S., Spyrou G., Ohno S.,
 RT "Activation of the JNK pathway by distantly related protein kinases,
 MEKK and MUK";
 RT Oncogene 12:641-650(1996).

DB 274 VAMMAPEVIRNEPSEKVIDWSPGVLMWELLTGEIPYKVDSSA--IIMVGNSNLSLHPV 331
QY 177 IKNLPRKPIESLMTRCWSKDPSPRSMEEIYKIMTHL 212
DB 332 PSSCPDGFKILLRQCWNKSPRRNPSFRQ--ILLHL 364

RESULT 8
MKC_MOUSE
ID MKC_MOUSE STANDARD; PRT; 888 AA.
AC Q60700; P70286;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Mitogen-activated protein kinase kinase 12 (EC 2.7.1.37)
DE (Leucine-zipper protein kinase) (ZPK) (Dual leucine zipper bearing kinase) (DLK).
GN MAP3K12 OR ZPK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CD-1; TISSUE=Brain;
RX MEDLINE=95074107; PubMed=7983011;
RA Holman L.B., Merrill S.E., Fan G.;
RT "Identification, molecular cloning, and characterization of dual leucine zipper bearing kinase. A novel serine/threonine protein kinase that defines a second subfamily of mixed lineage kinases.";
RT J. Biol. Chem. 269:30808-30817(1994);
RL [2]
RN SEQUENCE FROM N.A.
RC STRAIN=ICR X Swiss Webster; TISSUE=Brain;
RX MEDLINE=9635388; PubMed=8769565;
RA Blouin R., Beaudoin J., Bergeron P., Nadeau A., Gordin G.;
RT "Cell-specific expression of the ZPK gene in adult mouse tissues.";
RL DNA Cell Biol. 15:631-642(1996).
RN [3]
RN PHOSPHORYLATION, AND MUTAGENESIS OF LYS-185 AND GLU-192.
RP STRAIN=96279269; PubMed=8663324;
RX Meta M., Merritt S.E., Fan G., Yu G.G., Holman L.B.;
RT "Characterization of dual leucine zipper-bearing kinase, a mixed lineage kinase present in synaptic terminals whose phosphorylation state is regulated by membrane depolarization via calcineurin.";
RL J. Biol. Chem. 271:16888-16896(1996).
CC -1- FUNCTION: May be an activator of the JNK/SAPK pathway.
CC Phosphorylates beta-casein, histone 1 and myelin basic protein in vitro.
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- COFACTOR: Magnesium.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated.
CC -1- TISSUE SPECIFICITY: Expressed in brain, kidney, liver and pancreas, testis, gastrointestinal tract, stomach, liver and pancreas.
CC Within the nervous system, predominantly expressed in neurons and enriched in synaptic terminals.
CC -1- PTM: Autophosphorylated on Ser/Thr. Phosphorylated in cytosol under basal conditions and dephosphorylated when membrane-associated.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U14636; AAA57280.1; -
CC EMBL: U23789; AAB17123.1; -
CC HSPF; P12931; 1PMK.

DR MGD; MG1.1346881; Map3k12.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR004040; STY_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00221; STYKc1.
DR PROSITE; PS00107; PROTEIN KINASE ATP, FALSE_NEG.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
KW Transferase, Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation; Magnesium; Membrane.
FT DOMAIN 158 399
FT NP_BIND 164 172 ATP (BY SIMILARITY).
FT BINDING 185 185 ATP.
FT ACT_SITE 269 269 BY SIMILARITY.
FT DOMAIN 56 62 POLY-GLY.
FT DOMAIN 668 671 POLY-PRO.
FT DOMAIN 753 768 POLY-GLU.
FT MOTAGEN 185 185 K->A: NO CATALYTIC ACTIVITY.
FT MOTAGEN 192 192 E->A: NO CHANGE.
FT CONFLICT 18 18 V->A (IN REF. 2).
FT CONFLICT 28 29 KL->T (IN REF. 2).
FT CONFLICT 382 382 S->T (IN REF. 2).
FT CONFLICT 494 495 EQ->DE (IN REF. 2).
FT CONFLICT 517 517 N->D (IN REF. 2).
FT CONFLICT 794 794 E->G (IN REF. 2).
SQ SEQUENCE 888 AA; 96083 MW; CFCEFD134F899ABB CRC64;

Query Match 28.2%; Score 353.5; DB 1; Length 888;
Best Local Similarity 34.7%; Pred. No. 9,7e-26;
Matches 75; Conservative 43; Mismatches 81; Indels 17; Gaps 8;

QY 2 ELRQSRVNPVNVKLYGACLP--VCLWMEYEGSLYVNHGAPLPYTTAAHMSWC 59
DB 194 DIKRLRLKRPNIITPRGVCTQAPCYCIIMFPCAQOLYEVLRAGPV--TPSLLVDS 250
QY 60 LQSGGVAYLHSMQPKALIHRLDKPNNLLVAGTVLKI CDFGTACDIQTHMTNNK--GS 117
DB 251 MGIAAGMNVYHLRK--IIHRDLKSPN-MLITVDVVKISDFQSKELSDKSTKMSFAGT 306
QY 118 AAMMAPEVFEKSYSEKCVFSGIILMEVITRKRFDEIGAPFIMAV-HNGRRPL 176
DB 307 VAMMAPEVIRNEPSEKVIDWSPGVLMWELLTGEIPYKVDSSA--IIMVGNSNLSLHPV 364

QY 177 IKNLPRKPIESLMTRCWSKDPSPRSMEEIYKIMTHL 212
DB 365 PSSCPDGFKILLRQCWNKSPRRNPSFRQ--ILLHL 397

RESULT 9
KYK2_DICDI
ID KYK2_DICDI STANDARD; PRT; 410 AA.
AC P18161;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tyrosine-protein kinase 2 (EC 2.7.1.112) (Fragment).
GN PYKB OR DPYK2.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90287147; PubMed=1972546;
RA Tan J.L., Spudich J.A.;
RT "Developmentally regulated protein-tyrosine kinase genes in Dictyostelium discoideum.";
RL Mol. Cell. Biol. 10:3578-3583(1990).
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES.


```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M33784; AAA33203.1; --
CC PIR: B35670; B35670.
CC HSSP: P08631; IAD5.
CC DICTYDB: DD03011; pykB.
CC InterPro: IPR000719; Euk_pkinase.
CC InterPro: IPR004040; STY_pkinase.
CC InterPro: IPR001245; Tyr_pkinase.
CC Pfam: PF00069; pkinase; 1.
CC ProDom: PD000001; Euk_pkinase; 1.
CC SMART: SM00221; STYK; 1.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
CC PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
CC Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.
KW NON_TER
FT DOMAIN 108 381 PROTEIN_KINASE.
FT NP_BIND 114 122 ATP (BY SIMILARITY).
FT BINDING 135 135 ATP (BY SIMILARITY).
FT ACT_SITE 232 232 BY SIMILARITY.
SQ SEQUENCE 410 AA; 46386 MW; E9391B8605B9AEC1 CRC64;

Query Match 27.2%; Score 341; DB 1; Length 410;
Best Local Similarity 34.6%; Pred. No. 6.1e-25;
Matches 74; Conservative 42; Mismatches 84; Indels 14; Gaps 7;

QY 2 ELRQLSRVHPNIVKLYGACLPVCLVMEYAEAGSGSYLVHLH---GAEPUPYTAHAHMSW 58
DB 153 EVQNLKKGNHQNIVMPGICAYCPACITETYMAGSGSYLVHLHNSSTPKVYSFPLVLKM 212
QY 59 CLQCSQGVAYLHSMQPKALHRLDKPPNLLVAGGVVLKICDPGTACD--IQTHMTNNGK 116
DB 213 ATDMALGLHLHSHI---TIVHRDLTSQNLDELGNL-KISDFGLSAEKSRGSGMTMNG 268
QY 117 ---SAWMAPEVPEG-SNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMWAVHNGT 172
DB 269 GICNPWRPPELLTKNLGHYSEKVDVYCFSLVWVEIITGTIPESDLDG-SQSAQVAYAGL 327
QY 173 RPPLIKNLKPKIESLMTKRCWSDKPSRQSPWEEIV 206
DB 328 RPPPEYCDPELKLKLLTQCWEADPNDRPPPTIV 361

RESULT 10
CTR1_ARATH STANDARD; PRT; 821 AA.
AC Q05609;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Serine/threonine-protein kinase CTR1 (EC 2.7.1.37).
GN CTR1 OR AT5G03730 OR F17C15.150.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eumetazoa; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
ON [1]
RN SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia; TISSUE=Seedling;
RX MEDLINE=93161417; PubMed=8431946;
RA Kieber J.J., Rothenberg M., Roman G., Feldmann K.A., Ecker J.R.;
RA "CTR1, a negative regulator of the ethylene response pathway in
RT Arabidopsis, encodes a member of the raf family of protein kinases.";
RL Cell 72:427-441(1993).

```

```

RN RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016721; PubMed=11130714;
RA Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E.,
RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,
RA Kohara M., Matsumoto M., Matsumoto A., Muraki A., Nakayama S.,
RA Nakazaki N., Naruo K., Okumura S., Shingo S., Takeuchi C., Wada T.,
RA Watanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M.,
RA Huang E., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,
RA Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J.,
RA Stoneking T., Pepin K., Spieth J., Sekhon M., Armstrong J., Becker M.,
RA Belter E., Cordum H., Cordes M., Courtney L., Courtney W., Dante M.,
RA Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Latreille P.,
RA Leonard S., Meyer R., Mulvaney E., Ozersky P., Riley A., Strommatt C.,
RA Wagner-McPherson C., Wollam A., Yoakum M., Bell M., Dedhia N.,
RA Parnell L., Shah R., Rodriguez M., Hoon See L., Vil D., Baker J.,
RA Kirchoff K., Toth K., King L., Bahret A., Miller B., Maria M.,
RA Martienssen R., McCombie W.R., Wilson R.K., Murphy G., Bancroft I.,
RA Volckaert G., Wambutt R., Duesterhoeft A., Stiekema W., Pohl T.,
RA Entian K.-D., Terry N., Hartley N., Bent E., Johnson S.,
RA Langham S.-A., McCullagh B., Robben J., Grymonprez B., Zimmermann W.,
RA Ransperger U., Wedler H., Balke K., Wedler E., Peters S.,
RA van Staveren M., Dirkse W., Moolijman P., Klein Lankhorst R.,
RA Weitznegger T., Bothe G., Rose M., Hauf J., Berner S., Hempel S.,
RA Feldpausch M., Lamberth S., Villarroel R., Gielen J., Ardiles W.,
RA Bents O., Lemcke K., Kolesov G., Mayer K.F.X., Rudd S., Schoof H.,
RA Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Fransz P.F.;
RT "Sequence and analysis of chromosome 5 of the plant Arabidopsis
thaliana.";
RL Nature 408:823-826(2000).
CC -!- FUNCTION: ACTS AS A NEGATIVE REGULATOR IN THE ETHYLENE RESPONSE
PATHWAY.
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN BOTH SEEDLINGS AND ADULT PLANTS.
CC -!- MISCELLANEOUS: CTR1 MUTANTS DISPLAY ETHYLENE-TREATED PHENOTYPES,
CC RESULTING IN PLANTS WITH SMALL, UNEXPANDED LEAVES AND WHOSE SEED
CC COTYLEDON GROWTH IS IMPAIRED.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MTL/RAF SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L08789; AAA32779.1; --
CC EMBL: L08790; AAA32780.1; --
CC EMBL: AL162506; CAB82938.1; --
CC InterPro: IPR000719; Euk_pkinase.
CC InterPro: IPR004040; STY_pkinase.
CC InterPro: IPR002290; Ser_thr_pkinase.
CC Pfam: PF00069; pkinase; 1.
CC ProDom: PD000001; Euk_pkinase; 1.
CC SMART: SM00221; STYK; 1.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
CC Transferrase; Serine/threonine-protein kinase; ATP-binding.
KW DOMAIN 65 69 POLY-GLY.
FT DOMAIN 135 141 POLY-GLY.
FT DOMAIN 551 809 PROTEIN_KINASE.
FT NP_BIND 557 565 ATP (BY SIMILARITY).
FT BINDING 578 578 ATP (BY SIMILARITY).
FT ACT_SITE 676 676 BY SIMILARITY.
FT MUTAGEN 596 596 E->K: IN CTR1-4; EXHIBITS ETHYLENE-
FT TREATED PHENOTYPE.
FT MUTAGEN 694 694 D->E: IN CTR1-1; EXHIBITS ETHYLENE-
FT TREATED PHENOTYPE.
SQ SEQUENCE 821 AA; 90306 MW; 2922D3DCD0CC15BC CRC64;

```


RP SEQUENCE OF 1-280 FROM N.A.
 RX MEDLINE=89282215; PubMed=2660074;
 RA Takahashi M.;
 RT "Isolation of ret proto-oncogene cDNA with an amino-terminal signal
 sequence.";
 RL Oncogene 4:805-806(1989).
 RN [2]
 RP SEQUENCE OF 255-1114 FROM N.A.
 RX MEDLINE=90272230; PubMed=3078962;
 RA Takahashi M., Buma Y., Iwamoto T., Inaguma Y., Ikeda H., Hiai H.;
 RT "Cloning and expression of the ret proto-oncogene encoding a tyrosine
 kinase with two potential transmembrane domains.";
 RL Oncogene 3:571-578(1988).
 RN [3]
 RP SEQUENCE OF 588-1063 FROM N.A.
 RX MEDLINE=87257826; PubMed=3037315;
 RA Takahashi M., Cooper G.M.;
 RT "ret transforming gene encodes a fusion protein homologous to
 tyrosine kinases.";
 RL Mol. Cell. Biol. 7:1378-1385(1987).
 RN [4]
 RP REVIEW ON HSCR VARIANTS.
 RX MEDLINE=98023959; PubMed=9359036;
 RA Hofstra R.M.W., Osinga J., Buys C.H.C.M.;
 RT "Mutations in Hirschsprung disease: when does a mutation contribute to
 the phenotype.";
 RL Eur. J. Hum. Genet. 5:180-185(1997).
 RN [5]
 RP REVIEW ON VARIANTS.
 RX MEDLINE=97220587; PubMed=9067749;
 RA Eng C., Mulligan L.M.;
 RT "Mutations of the RET proto-oncogene in the multiple endocrine
 neoplasia type 2 syndromes, related sporadic tumours, and
 hirschsprung disease.";
 RL Hum. Mutat. 9:97-109(1997).
 RN [6]
 RP VARIANTS MEN2A/MTC W-611; S-618; R-620; Y-620 AND R-634.
 RX MEDLINE=93372843; PubMed=8103403;
 RA Donis-Keller H., Dou S., Chi D., Carlson K.M., Toshima K.,
 RA Laifmore T.C., Howe J.R., Moley J.F., Goodfellow P., Wellis S.A. Jr.;
 RT "Mutations in the RET proto-oncogene are associated with MEN 2A and
 FMTC.";
 RL Hum. Mol. Genet. 2:851-856(1993).
 RN [7]
 RP VARIANTS G-618; 632-DVR-634; G-634; F-634; Y-634 AND S-634.
 RX MEDLINE=93275414; PubMed=8099202;
 RA Mulligan L.M., Kwok J.B.J., Healey C.S., Elsdon M.J., Eng C.,
 RA Gardner E., Love D.R., Mole S.E., Moore J.K., Papi L., Ponder M.A.,
 RA Telenius H., Tunnacliffe A., Ponder B.A.J.;
 RT "Germ-line mutations of the RET proto-oncogene in multiple endocrine
 neoplasia type 2A.";
 RL Nature 363:458-460(1993).
 RN [8]
 RP VARIANTS HSCR P-40; L-399; Q-762; P-765; Q-897; G-972 AND L-973.
 RX MEDLINE=95219414; PubMed=7704557;
 RA Yin L., Barone V., Sri M., Bolino A., Bocciardi R., Ceccherini I.,
 RA Fasini B., Tocco T., Lerone M., Cywes S., Moore S.,
 RA Vanderwinden J.-M., Abramowicz M.J., Kristoffersson U., Larsson L.T.,
 RA Hamel B.C.J., Silengo M., Martucciello G., Romeo G.;
 RT "Heterogeneity and low detection rate of RET mutations in Hirschsprung
 disease.";
 RL Eur. J. Hum. Genet. 2:272-280(1994).
 RN [9]
 RP VARIANT MEN2B THR-918.
 RX MEDLINE=94272459; PubMed=7911697;
 RA Eng C., Smith D.P., Mulligan L.M., Nagai M.A., Healey C.S.,
 RA Ponder M.A., Gardner E., Scheumann G.F., Jackson C.E., Tunnacliffe A.,
 RA Ponder B.A.J.;
 RT "Point mutation within the tyrosine kinase domain of the RET
 proto-oncogene in multiple endocrine neoplasia type 2B and related
 sporadic tumours.";
 RL Hum. Mol. Genet. 3:237-241(1994).
 RN [10]

RP VARIANTS FMTC/MEN2A R-618; S-618; F-620; R-620; F-634; G-634 AND
 Y-634.
 RX MEDLINE=94348513; PubMed=7915165;
 RA Xue F., Yu H., Maurer L.H., Menoli V.A., Nutille-Mcmenemy N.,
 RA Schuster M.K., Browden D.W., Mao J.-I., Noll W.W.;
 RT "Germline RET mutations in MEN 2A and FMTC and their detection by
 simple DNA diagnostic tests.";
 RL Hum. Mol. Genet. 3:635-638(1994).
 RN [11]
 RP VARIANTS FMTC/MEN2A TYR-609; ARG-618; SER-618 AND SER-620.
 RX MEDLINE=95152521; PubMed=7849720;
 RA Blaugrund J.E., Johns M.M. Jr., Eby Y.J., Ball D.W., Baylin S.B.,
 RA Ruban R.H., Sidransky D.;
 RT "RET proto-oncogene mutations in inherited and sporadic medullary
 thyroid cancer.";
 RL Hum. Mol. Genet. 3:1895-1897(1994).
 RN [12]
 RP VARIANTS FMTC, AND VARIANTS MEN2A.
 RX MEDLINE=95179108; PubMed=7874109;
 RA Schuffenecker I., Billaud M., Calender A., Chambe B., Ginot N.,
 RA Calmettes C., Modigliani E., Lenoir G.M.;
 RT "RET proto-oncogene mutations in French MEN 2A and FMTC families.";
 RL Hum. Mol. Genet. 3:1939-1943(1994).
 RN [13]
 RP VARIANT HSCR W-609, VARIANT HSCR/MEN2A R-618, AND VARIANT HSCR/FMTC
 R-620.
 RX MEDLINE=95187155; PubMed=7881414;
 RA Mulligan L.M., Eng C., Attie T., Lyonnet S., Marsh D.J., Hyland V.J.,
 RA Robinson B.G., Frilling A., Verellen-Dumoulin C., Safar A.,
 RA Venter D.J., Munnich A., Ponder B.A.J.;
 RT "Diverse phenotypes associated with exon 10 mutations of the RET
 proto-oncogene.";
 RL Hum. Mol. Genet. 3:2163-2167(1994).
 RN [14]
 RP VARIANT MEN2B THR-918.
 RX MEDLINE=94159102; PubMed=7906866;
 RA Hofstra R.M.W., Landsvater R.M., Ceccherini I., Stulp R.P.,
 RA Stelwagen T., Luo Y., Pasini B., Hoepfner J.W.M., van Amstel H.K.P.,
 RA Romeo G., Lips C.J.M., Buys C.H.C.M.;
 RT "A mutation in the RET proto-oncogene associated with multiple
 endocrine neoplasia type 2B and sporadic medullary thyroid
 carcinoma.";
 RL Nature 367:375-376(1994).
 RN [15]
 RP VARIANTS HSCR PRO-765; GLN-897 AND GLY-972.
 RX MEDLINE=94159103; PubMed=8114938;
 RA Romeo G., Ronchetto P., Luo Y., Barone V., Seri M., Ceccherini I.,
 RA Pasini B., Bocciardi R., Lerone M., Kaarlainen H., Martucciello G.;
 RT "Point mutations affecting the tyrosine kinase domain of the RET
 proto-oncogene in Hirschsprung's disease.";
 RL Nature 367:377-378(1994).
 RN [16]
 RP VARIANTS HSCR LEU-32; LEU-64; GLN-330 AND LEU-393.
 RX MEDLINE=94159104; PubMed=8114939;
 RA Edery P., Lyonnet S., Mulligan L.M., Pelet A., Dow E., Abel L.,
 RA Holder S., Nihoul-Fkete C., Ponder B.A.J., Munnich A.;
 RT "Mutations of the RET proto-oncogene in Hirschsprung's disease.";
 RL Nature 367:378-380(1994).
 RN [17]
 RP VARIANT MEN2B THR-918.
 RX MEDLINE=94151373; PubMed=7906417;
 RA Carlson K.M., Dou S., Chi D., Scavarda N., Toshima K., Jackson C.E.,
 RA Wells S.A. Jr., Goodfellow P.J., Donis-Keller H.;
 RT "Single missense mutation in the tyrosine kinase catalytic domain of
 the RET protooncogene is associated with multiple endocrine neoplasia
 type 2B.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:1579-1583(1994).
 RN [18]
 RP VARIANTS MTC; FMTC; MEN2A AND MEN2B.
 RX MEDLINE=96223053; PubMed=8625130;
 RA Komminoth P., Kunz E.K., Matias-Guiu X., Hiort O., Christensen G.,
 RA Colomer A., Roth J., Heitz P.U.;
 RT "Analysis of RET protooncogene point mutations distinguishes heritable

RT from nonheritable medullary thyroid carcinomas.";
 RL Cancer 76:479-489(1995).
 RN [19]
 RP VARIANTS MENZA SER-618; SER-620; ARG-634 AND TYR-634.
 RX MEDLINE=95163936; PubMed=7860065;
 RA Takiguchi-Shirahama S., Koyama K., Mitsuuchi A., Wakasugi T., Oishi S.,
 RA Takami H., Hikiji K., Nakamura Y.;
 RT "Germline mutations of the RET proto-oncogene in eight Japanese
 RT patients with multiple endocrine neoplasia type 2A (MEN2A).";
 RL Hum. Genet. 95:187-190(1995).
 RN [20]
 RP VARIANTS HSCR L-20; S-93; Q-330; Y-609 AND R-620, AND VARIANT C-982.
 RC TISSUE=BLOOD;
 RX MEDLINE=95360000; PubMed=7633441;
 RA Anglist M., Bolik S., Thiel B., Puffenberger E.G., Hofstra R.M.W.,
 RA Buys C.H.C.M., Cass D.T., Chakravarti A.;
 RT "Mutation analysis of the RET receptor tyrosine kinase in Hirschsprung
 RT disease.";
 RL Hum. Mol. Genet. 4:821-830(1995).
 RN [21]
 RP VARIANTS HSCR.
 RC TISSUE=Leukocyte;
 RX MEDLINE=96090258; PubMed=7581377;
 RA Attie T., Pelet A., Ederly P., Eng C., Mulligan L.M., Amiel J.,
 RA Bourard L., Beldjord C., Nihoul-Fekete C., Munnich A., Ponder B.A.J.,
 RA Lyonnet S.;
 RT "Diversity of RET proto-oncogene mutations in familial and sporadic
 RT Hirschsprung disease.";
 RL Hum. Mol. Genet. 4:1381-1386(1995).
 RN [22]
 RP VARIANT MEN2B THR-918, AND VARIANT TYR-922.
 RX MEDLINE=96121602; PubMed=8595427;
 RA Kitemura Y., Scavarda N., Wells S.A. Jr., Jackson C.E.,
 RA Goodfellow P.J.;
 RT "Two maternally derived missense mutations in the tyrosine kinase
 RT domain of the RET protooncogene in a patient with de novo MEN 2B.";
 RL Hum. Mol. Genet. 4:1987-1988(1995).
 RN [23]
 RP VARIANT FMTC ASP-768.
 RX MEDLINE=95148221; PubMed=7845675;
 RA Eng C., Smith D.P., Mulligan L.M., Healey C.S., Zvelebil M.J.,
 RA

Query Match 24.3%; Score 304.5; DB 1; Length 1114;
 Best Local Similarity 32.1%; Pred. No. 5.6e-21;
 Matches 79; Conservative 44; Mismatches 78; Indels 45; Gaps 9;

OY 2 ELRQDSRVNHPNIVKLYGACT--NPFCLVMEYAEGLSYLVNHLGAEPV-PYY----- 50
 DB 775 EFNVLKQVNHPRVIVKLYGACSGDGPLLLIVEYAKYGLRGFLRESRKVGPGLSGGSRN 834
 OY 51 -----TAHMSWCIQCSQGVAYLHSMQPKALIRHDLKPPVLLVAGTVALKI 98
 DB 835 SSSLDHPERALTWGDLISFAMWISQGMQYLAEMK--LVHRDLAARN-ILVAGGRKKKI 890
 OY 99 CDFGTACDI---QTMNTNKG--SAAMAPVEFGESNYSEKCDVFSWIIIMEVITRRKP 153
 DB 831 SDFGSRVYEDSDYVKSQGRIPVKMAISLPHVITTSQDWSRGVLMELVI----- 946
 OY 154 FDEIGG-----PAFRIMAVANGTRPPLIKNLPKPIESLMTRCMSKDPQSPRSMEEIVK 207
 DB 947 ---LOGNFPYPGIPRPLNLLKTGHRMERPDNCSSEMYRLMLQCKWQPPDRKRVFADISK 1003
 OY 208 IMTHLM 213
 DB 1004 DLEKMM 1009

RESULT 13
 TEC_MOUSE STANDARD; PRT; 630 AA.
 ID TEC_MOUSE
 AC P24604;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tyrosine-protein kinase Tec (EC 2.7.1.112).
 GN TEC.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=93149603; PubMed=7678927;
 RA Mano H., Mano K., Tang B., Koehler M., Yi T., Gilbert D.J.,
 RA Jenkins N.A., Copeland N.G., Ihle J.N.;
 RT "Expression of a novel form of Tec kinase in hematopoietic cells and
 RT mapping of the gene to chromosome 5 near Kit.";
 RL Oncogene 8:417-424(1993).
 RN [2]
 RP SEQUENCE OF 82-630 FROM N.A. (ISOFORM 2).
 RC STRAIN=BALB/C; TISSUE=Liver.
 RX MEDLINE=91133729; PubMed=2284097;
 RA Mano H., Ishikawa F., Nishida J., Hirai H., Takaku F.;
 RT "A novel protein-tyrosine kinase, tec, is preferentially expressed in
 RT liver.";
 RL Oncogene 5:1781-1786(1990).
 RN [3]
 RP SEQUENCE OF 485-553 FROM N.A.
 RX MEDLINE=90152381; PubMed=2482828;
 RA Wilks A.F., Kurban R.R., Hovense C.M., Ralph S.J.;
 RT "The application of the polymerase chain reaction to cloning members
 RT of the protein tyrosine kinase family.";
 RL Gene 85:67-74(1989).
 CC -!- FUNCTION: MAY BE AN IMPORTANT SIGNAL TRANSDUCER FOR CELL DIVISION
 CC AND/OR FOR DIFFERENTIATION IN THE LIVER SYSTEM.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -!- ALTERNATIVE PRODUCTS: 3 isoforms; 1 (shown here) and 2; are
 CC produced by alternative splicing. Isoform 3 may be very rare or
 CC due to a cloning artifact.
 CC -!- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN LIVER. EXPRESSION
 CC IS ALSO SEEN IN THE HEMATOPOIETIC CELLS SUCH AS BONE MARROW,
 CC THYMUS AND SPLEEN. LOWER EXPRESSION SEEN IN THE HEART, KIDNEY AND
 CC OVARY.
 CC -!- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. BTK
 CC SUBFAMILY.
 CC -!- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -!- SIMILARITY: CONTAINS 1 PH DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: S53716; AAA13515.2; -.
 CC EMBL: X55663; CAA39196.1; -.
 CC EMBL: M33427; AAA40018.1; -.
 CC PIR: S13763; S13763.
 CC HSP: 006187; 1855.
 CC MGD: MGI:98662; Tec.
 CC InterPro: IPR001562; BTK.
 CC InterPro: IPR000719; Euk_pkinase.
 CC InterPro: IPR001849; SH2.
 CC InterPro: IPR000980; SH2.
 CC InterPro: IPR001452; SH3.
 CC InterPro: IPR001245; Tyr_pkinase.
 CC Pfam: PF00017; SH2; 1.
 CC Pfam: PF00018; SH3; 1.
 CC Pfam: PF00069; SH3; 1.
 CC Pfam: PF00169; PH; 1.
 CC Pfam: PF00779; BTK; 1.


```
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00779; BTK; 1.
DR PRINTS; PR00401; SH2DOMAIN.
DR PRINTS; PR00402; TECBTKDOMAIN.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Euk_pkinase; 1.
DR ProDom; PD000086; SH3; 1.
DR ProDom; PD000093; SH2; 1.
DR SMART; SM00107; BTK; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00252; SH2; 1.
DR SMART; SM00326; SH3; 1.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS50001; SH2; 1.
DR PROSITE; PS50002; SH3; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
KW Transferase; Tyrosine-protein kinase; Phosphorylation;
KW ATP-binding; SH2 domain; SH3 domain; 3D-structure.
FT DOMAIN 4 117 PH.
FT DOMAIN 177 237 SH3.
FT DOMAIN 245 343 SH2.
FT DOMAIN 368 620 PROTEIN KINASE.
FT NP_BIND 374 382 ATP (BY SIMILARITY).
FT BINDING 396 396 ATP (BY SIMILARITY).
FT ACT_SITE 487 487 BY SIMILARITY.
FT MOD_RES 517 517 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CONFLICT 82 87 MISSING (IN REF. 2, 3 AND 4).
FT CONFLICT 535 535 F -> S (IN REF. 3).
FT CONFLICT 540 540 Y -> C (IN REF. 3).
SQ SEQUENCE 625 AA; 72291 MW; F7A4A18A8A1AADD C RC64;

Query Match 23.9%; Score 299.5; DB 1; Length 625;
Best Local Similarity 32.1%; Pred. No. 8.6e-21;
Matches 72; Conservative 37; Mismatches 88; Indels 27; Gaps 9;

QY 2 ELROLSRVNHPNIVKLYGACLN--PVCVMVEYAEQGGSLYNVLHGAEPLPYVTTAAHAMSWC 59
Db 411 EAEVMMKLSHPKLVQLYGVCLQAPICLVFEPWEHGCGLDYLRSQRL--FAAETLLGMC 468
QY 60 LQCSQGVAYLHSMQPKALIHRLKPPNLLVAGGTVLKICDPG----TACDIQTHMTNNK 115
Db 469 LDVCEGMAYL--EKACVIRHDLAARN-CLVGENQVIKVSDFGMTRFVLDDQVTSSTGTK 524
QY 116 GSAAWMAPEVPEGSNYSEKCDVPSWGIILWEVITRRK-PFDEIGGPAFRIMWAVHNGTRP 174
Db 525 FPKWASPEVFSFSGYSKSDVMSFGVLMWEVFESEKIPYENRSNS--EVVEDISTGFR- 581
QY 175 PLIKNLPKP-----IESLWTRCWSKDPSPORSPSMEEIVKIMTHL 212
Db 582 -----LYKPLASCHVYQIMNHCWKKEPEDRPPFPQSLLSLAEI 620
```

Search completed: December 10, 2002, 03:57:41
Job time : 18 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 10, 2002, 03:51:50 ; Search time 31 Seconds
(without alignments)

1515.443 Million cell updates/sec

Title: US-09-830-144-2_COPY_76_303

Perfect score: 1252

Sequence: 1 VELQLSRVNHFNIVKLYGA.....MTHLMRYFFGADPLQLYPCQ 228

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.podent:*
12: sp.podent:*
13: sp.podent:*
14: sp.podent:*
15: sp.podent:*
16: sp.podent:*
17: sp.podent:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1252	100.0	478	4 Q9NTR4	Q9ntr4 homo sapien
2	1252	100.0	491	4 Q9NZ70	Q9nz70 homo sapien
3	1252	100.0	539	4 Q9NTR1	Q9ntr1 homo sapien
4	1252	100.0	566	4 Q9NTR2	Q9ntr2 homo sapien
5	1252	100.0	606	11 Q923A8	Q923a8 mus musculu
6	1221	97.5	616	13 Q923A8	Q923a8 mus musculu
7	700	55.9	678	5 Q9V3Q6	Q9v3q6 drosophila
8	414	33.1	252	5 Q9VCV0	Q9vcv0 drosophila
9	376	30.0	371	13 Q9QZY8	Q9qzy8 brachydanio
10	374	29.9	411	10 Q9ZQ31	Q9zq31 arabidopsis
11	373	29.8	412	10 Q9M085	Q9m085 arabidopsis
12	372	29.7	454	11 Q9ESL3	Q9esl3 mus musculu
13	372	29.7	802	11 Q9ESL4	Q9esl4 mus musculu
14	371	29.6	455	4 Q9HCC4	Q9hcc4 homo sapien
15	371	29.6	800	4 Q9HDD2	Q9hdd2 homo sapien
16	371	29.6	800	4 Q9HCC5	Q9hcc5 homo sapien

17	371	29.6	800	4 Q9NYL2	Q9nyl2 homo sapien
18	371	29.6	800	4 Q9NYB9	Q9nyb9 homo sapien
19	368	29.4	416	10 Q94C42	Q94c42 triticum ae
20	366	29.2	637	10 Q94J41	Q94j41 oryza sativ
21	363.5	29.0	1020	5 Q9W313	Q9w313 drosophila
22	363.5	29.0	1148	5 Q9SVF6	Q9svf6 drosophila
23	363.5	29.0	1161	5 Q9SUN8	Q9sun8 drosophila
24	362	28.9	422	5 Q23846	Q23846 dictyosteli
25	362	28.9	1338	5 Q23927	Q23927 dictyosteli
26	360.5	28.8	1066	4 Q9H2N5	Q9h2n5 homo sapien
27	356	28.4	462	10 Q98866	Q98866 glycine max
28	355.5	28.4	370	10 Q9S7D5	Q9s7d5 arabidopsis
29	355	28.4	564	4 Q9H1Y7	Q9h1y7 homo sapien
30	353.5	28.2	859	4 Q8WY25	Q8wy25 homo sapien
31	351.5	28.1	977	5 Q9VW24	Q9vw24 drosophila
32	351	28.0	570	4 Q8WVW2	Q8wv22 homo sapien
33	351	28.0	1036	4 Q8WVW1	Q8wv21 homo sapien
34	350	28.0	546	10 Q22558	Q22558 arabidopsis
35	349	27.9	1001	11 Q8VDG6	Q8vdg6 mus musculu
36	348	27.8	391	10 Q22100	Q22100 arabidopsis
37	347.5	27.8	850	11 Q9DJ15	Q9dj15 mus musculu
38	347	27.7	346	10 Q9FGS7	Q9fgs7 arabidopsis
39	347	27.7	475	10 Q9STG5	Q9stg5 arabidopsis
40	346	27.6	570	10 Q8RWL6	Q8rw16 arabidopsis
41	346	27.6	740	5 Q21982	Q21982 caenorhabdi
42	346	27.6	1030	10 Q9C9U5	Q9c9u5 arabidopsis
43	345.5	27.6	553	10 Q81808	Q81808 arabidopsis
44	345.5	27.6	847	4 Q16584	Q16584 homo sapien
45	341.5	27.3	405	10 Q9LVQ9	Q9lvq9 arabidopsis

ALIGNMENTS

RESULT 1

Q9NTR4 ID Q9NTR4 PRELIMINARY; PRT; 478 AA.
AC Q9NTR4; 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 20, Last annotation update)
DE DJ154G14.1.3 (Mitogen-activated protein kinase kinase 7
DE (TGF-beta activated kinase 1c (TAK1))) (Fragment).
GN MAP3K7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tracey A.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AL121964; CAB87604.1; -.
DR HSSP; P12931; 1FMK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR004040; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00221; STYKC; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
FT Kinase, Serine/threonine-protein kinase.
FT NON TER
SQ SEQUENCE 478 AA; 52482 MW; 177CC8CFA8D8DBF8 CRC64;

Query Match 100.0%; Score 1252; DB 4; Length 478;
Best Local Similarity 100.0%; Pred. No. 1.5e-118;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 VELRLSRVNHNPNI VKLYGACLNPCVLWMEYAEAGSLYNNVHGAEPPLPYTTAAHMSWCL 60
DB 36 VELRLSRVNHNPNI VKLYGACLNPCVLWMEYAEAGSLYNNVHGAEPPLPYTTAAHMSWCL 95
QY 61 QCSGVAAYLHSMQPKALIHRLDKPNNLLVAGTGLKICDFTGACDIQTMTNKKSAAM 120
DB 96 QCSGVAAYLHSMQPKALIHRLDKPNNLLVAGTGLKICDFTGACDIQTMTNKKSAAM 155
QY 121 MAPEVEGSSNYSEKCDVFSWGIIIMEVITRRKPFDEIGGPAFRIMAVHNGTRPPLIKNL 180
DB 156 MAPEVEGSSNYSEKCDVFSWGIIIMEVITRRKPFDEIGGPAFRIMAVHNGTRPPLIKNL 215
QY 181 PKPIESLMTRCMSKDPSPSPSMEIIVKIMTHLMRYFPGADEPLQYPCQ 228
DB 216 PKPIESLMTRCMSKDPSPSPSMEIIVKIMTHLMRYFPGADEPLQYPCQ 263

RESULT 2
Q9NZ70 PRELIMINARY; PRT; 491 AA.
ID Q9NZ70; Q9NZ70;
AC Q9NZ70; Q9NZ70;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE TGF-beta-activated kinase splice variant d (DJ154G14.1.4)
DE (mitogen-activated protein kinase kinase kinase 7 (TGF-beta activated kinase 1d (TAK1)))
GN TAK1 OR MAP3K7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20568765; PubMed=11118615;
RA Dempsey C.E., Sakurai H., Sugita T., Guesdon F.;
RT "Alternative splicing and gene structure of the transforming growth factor beta-activated kinase 1."
RT Biochim. Biophys. Acta 1517:46-52 (2000).
RN [2]
RP SEQUENCE OF 41-491 FROM N.A.
RA Tracey A.;
RC Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: AF218074; AAF27652.1; -.
DR EMBL: AL121964; CAB87605.1; -.
DR HSSP: P08631; 1ADM.
DR HSSP: P12931; 1PMK.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR004040; STY_pkinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00069; pkinase.1.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00221; STYKc; 1.
DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE: PS0011; PROTEIN KINASE DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferrase.
SQ SEQUENCE 491 AA; 53739 MW; B7D8832E286A99C5 CRC64;
```

```
Query Match 100.0%; Score 1252; DB 4; Length 491;
Best Local Similarity 100.0%; Pred. No. 1.6e-118;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 VELRLSRVNHNPNI VKLYGACLNPCVLWMEYAEAGSLYNNVHGAEPPLPYTTAAHMSWCL 60
DB 76 VELRLSRVNHNPNI VKLYGACLNPCVLWMEYAEAGSLYNNVHGAEPPLPYTTAAHMSWCL 135
QY 61 QCSGVAAYLHSMQPKALIHRLDKPNNLLVAGTGLKICDFTGACDIQTMTNKKSAAM 120
```

```
DB 136 QCSGVAAYLHSMQPKALIHRLDKPNNLLVAGTGLKICDFTGACDIQTMTNKKSAAM 195
QY 121 MAPEVEGSSNYSEKCDVFSWGIIIMEVITRRKPFDEIGGPAFRIMAVHNGTRPPLIKNL 180
DB 196 MAPEVEGSSNYSEKCDVFSWGIIIMEVITRRKPFDEIGGPAFRIMAVHNGTRPPLIKNL 255
QY 181 PKPIESLMTRCMSKDPSPSPSMEIIVKIMTHLMRYFPGADEPLQYPCQ 228
DB 256 PKPIESLMTRCMSKDPSPSPSMEIIVKIMTHLMRYFPGADEPLQYPCQ 303
```

```
RESULT 3
Q9NTR1 PRELIMINARY; PRT; 539 AA.
ID Q9NTR1; Q9NTR1;
AC Q9NTR1; Q9NTR1;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE DJ154G14.1.1 (mitogen-activated protein kinase kinase kinase 7 (TGF-beta activated kinase 1a (TAK1))) (Fragment).
GN MAP3K7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tracey A.;
RC Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: AL121964; CAB87607.1; -.
DR HSSP: P12931; 1PMK.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR004040; STY_pkinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00069; pkinase.1.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00221; STYKc; 1.
DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE: PS0011; PROTEIN KINASE DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW Kinase; Serine/threonine-protein kinase.
FT NON_TER 1
SQ SEQUENCE 539 AA; 60006 MW; B6183F53CC7F324 CRC64;
```

```
Query Match 100.0%; Score 1252; DB 4; Length 539;
Best Local Similarity 100.0%; Pred. No. 1.7e-118;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 VELRLSRVNHNPNI VKLYGACLNPCVLWMEYAEAGSLYNNVHGAEPPLPYTTAAHMSWCL 60
DB 36 VELRLSRVNHNPNI VKLYGACLNPCVLWMEYAEAGSLYNNVHGAEPPLPYTTAAHMSWCL 95
QY 61 QCSGVAAYLHSMQPKALIHRLDKPNNLLVAGTGLKICDFTGACDIQTMTNKKSAAM 120
DB 96 QCSGVAAYLHSMQPKALIHRLDKPNNLLVAGTGLKICDFTGACDIQTMTNKKSAAM 155
QY 121 MAPEVEGSSNYSEKCDVFSWGIIIMEVITRRKPFDEIGGPAFRIMAVHNGTRPPLIKNL 180
DB 156 MAPEVEGSSNYSEKCDVFSWGIIIMEVITRRKPFDEIGGPAFRIMAVHNGTRPPLIKNL 215
QY 181 PKPIESLMTRCMSKDPSPSPSMEIIVKIMTHLMRYFPGADEPLQYPCQ 228
DB 216 PKPIESLMTRCMSKDPSPSPSMEIIVKIMTHLMRYFPGADEPLQYPCQ 263
```

```
RESULT 4
Q9NTR2 PRELIMINARY; PRT; 566 AA.
ID Q9NTR2; Q9NTR2;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
```

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE DJ154G14.1.2 (Mitogen-activated protein kinase kinase 7
DE (TGF-beta activated kinase 1b (TAK1))) (Fragment).
GN MAP3K7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tracey A.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBSJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC EMBL; ALI21964; CAB87606.1; -.
DR HSSP; PI2931; 1FMK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR004040; STY_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00221; STYKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Kinase; Serine/threonine-protein_kinase.
FT NON_TER 1
SQ SEQUENCE 566 AA; 62972 MW; 32CDAC1211B200CF CRC64;
Query Match 100.0%; Score 1252; DB 4; Length 566;
Best Local Similarity 100.0%; Pred. No. 1.8e-118;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VELQLSRVNHNPVIVKLYGACLPVCLVMEYAEAGGSLYNVLHGAEPPLPYTTAAHAMSWCL 60
DB 36 VELQLSRVNHNPVIVKLYGACLPVCLVMEYAEAGGSLYNVLHGAEPPLPYTTAAHAMSWCL 95
QY 61 QCSQGVAYLHSMOPKALIHRLDLPNLLVAGGTVLKICDFGTACDIQTHMTNNKGSAAW 120
DB 96 QCSQGVAYLHSMOPKALIHRLDLPNLLVAGGTVLKICDFGTACDIQTHMTNNKGSAAW 155
QY 121 MAPEVFGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 180
DB 156 MAPEVFGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 215
QY 181 PKPIESLMTRCWSKDPQRSMEIIVKIMTHLMRYFPFGADEPLQYPCQ 228
DB 216 PKPIESLMTRCWSKDPQRSMEIIVKIMTHLMRYFPFGADEPLQYPCQ 263
RESULT 5
Q923A8 ID Q923A8 PRELIMINARY; PRT; 606 AA.
AC Q923A8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Unknown (protein for MGC:5989).
GN MAP3K7.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBSJ databases.
DR EMBL; BC006665; AAH06665.1; -.
DR MGD; MGI:1346877; Map3k7.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.

DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN 1.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; UNKNOWN 1.
KW ATP-binding; Transferase.
SQ SEQUENCE 606 AA; 67194 MW; AB8664F389272102 CRC64;
Query Match 100.0%; Score 1252; DB 11; Length 606;
Best Local Similarity 100.0%; Pred. No. 2e-118;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VELQLSRVNHNPVIVKLYGACLPVCLVMEYAEAGGSLYNVLHGAEPPLPYTTAAHAMSWCL 60
DB 76 VELQLSRVNHNPVIVKLYGACLPVCLVMEYAEAGGSLYNVLHGAEPPLPYTTAAHAMSWCL 135
QY 61 QCSQGVAYLHSMOPKALIHRLDLPNLLVAGGTVLKICDFGTACDIQTHMTNNKGSAAW 120
DB 136 QCSQGVAYLHSMOPKALIHRLDLPNLLVAGGTVLKICDFGTACDIQTHMTNNKGSAAW 195
QY 121 MAPEVFGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 180
DB 196 MAPEVFGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 255
QY 181 PKPIESLMTRCWSKDPQRSMEIIVKIMTHLMRYFPFGADEPLQYPCQ 228
DB 256 PKPIESLMTRCWSKDPQRSMEIIVKIMTHLMRYFPFGADEPLQYPCQ 303
RESULT 6
O73613 ID O73613 PRELIMINARY; PRT; 616 AA.
AC O73613;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE TAK1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98130593; PubMed=9463380;
RA Shibuya H., Iwata H., Masuyama N., Gotoh Y., Yamaguchi K., Irie K.,
RA Matsumoto K., Nishida E., Ueno N.;
RT "Role of TAK1 and TAB1 in BMP signaling in early Xenopus
RT development.";
RL EMBO J. 17:1019-1028(1998).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; U92030; AAC14008.1; -.
DR HSSP; PI2931; 1FMK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR004040; STY_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00221; STYKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Serine/threonine-protein_kinase; Transferase.
SQ SEQUENCE 616 AA; 68464 MW; 493AD2A05ADC38B6 CRC64;
Query Match 97.5%; Score 1221; DB 13; Length 616;
Best Local Similarity 97.4%; Pred. No. 2.8e-115;
Matches 222; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 VELQLSRVNHNPVIVKLYGACLPVCLVMEYAEAGGSLYNVLHGAEPPLPYTTAAHAMSWCL 60
DB 65 VELQLSRVNHNPVIVKLYGACLPVCLVMEYAEAGGSLYNVLHGAEPPLPYTTAAHAMSWCL 124

```
QY 61 QCSGVAYLHSMQPKALIHRLDKPRLLLVAGTVLKICDFGTACDIOTHTMNNKSSAA 120
DB 125 QCAQGAVALHSMKPKALIHRLDKPRLLLVAGTVLKICDFGTACDIOTHTMNNKSSAA 184
QY 121 MAEEVEGNSYSEKCDVFSWGIIILMEVITRRKPFDEIGGAPRIMAAVANGTRPPLIKL 180
DB 185 MAEEVEGNSYSEKCDVFSWGIIILMEVITRRKPFDEIGGAPRIMAAVANGTRPPLIKL 244
QY 181 PKPIESIMTRCWSKDSQRPSPMEIYKIMTHLMRYFGADEPLQYPCQ 228
DB 245 PKPIESIMTRCWSKDSQRPSPMEIYKIMTHLMRYFGADEPLQYPCQ 292

RESULT 7
Q9V306 PRELIMINARY; PRT; 678 AA.
ID Q9V306;
AC Q9V306;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Putative TAK1 protein (CG1388 protein) (LD42274P).
GN TAK1 OR CG1388 OR CG18492.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Ephydroidea; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxId=7227;
RX STRAIN=BERKELEY;
RC MEDLINE=20196006; PubMed=10731132;
RA ADAMS M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandal D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Burks K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos E., Delcher A.L., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Jimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skuski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock K.C., Wu D., Yang S., Yao Q.A.,
RA Williams S.M., Woodage T., Worley K.G., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang Q., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Takatsu Y., Nakamura M., Stapleton M., Danos M., Matsumoto M.,
RA O'Connor M.B., Shibuya H., Ueno N.;
```

```
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carleon J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise B., George R.,
RA Gonzalez M., Garin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacle J., Paragae V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.B., Rubin G.M., Celniker S.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -1 SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: AE003571; AAF50895.1; -
DR EMBL: AF199466; AAF06815.1; -
DR EMBL: AY051953; AAK93377.1; -
DR HSSP: P08631; 1AD5.
DR FlyBase: FBgn0026323; Tak1.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR001990; Grabin.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR004040; STY_pkinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00221; STYKC; 1.
DR PROSITE: PS00422; GRANINS_1; UNKNOWN_1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR ATP-binding; Hypothetical protein; Kinase;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 678 AA; 75674 MW; 87EBA80CDBBCD45 CRC64;

Query Match 55.3%; Score 700; DB 5; Length 678;
Best Local Similarity 57.5%; Pred. No. 2,2e-62;
Matches 130; Conservative 35; Mismatches 57; Indels 4; Gaps 3;

QY 2 ELRQLSRNHNHYLYKG--ACLNPCVLNMEYAGGSLVYVHGAEPLPYTAAHMSWC 59
DB 60 EVKQLSRKXHNPIIALIGSSYQCATYLIWFEGGSLHNFLLHG-KVXPAYSLAHMSWA 118
QY 60 LQCSGVAYLHSMQPKALIHRLDKPRLLLVAGTVLKICDFGTACDIOTHTMNNKSSAA 119
DB 119 RQCBAGLALYHAMPKPLIHRLDKPRLLLVAGTVLKICDFGTADKSTMTMNNKSSAA 178
QY 120 MAEEVEGNSYSEKCDVFSWGIIILMEVITRRKPFDEIGGAPRIMAAVANGTRPPLIKL 179
DB 179 MAEEVEGNSYSEKCDVFSWGIIILMEVITRRKPFDEIGGAPRIMAAVANGTRPPLIKL 237
QY 180 PKPIESIMTRCWSKDSQRPSPMEIYKIMTHLMRYFGADEPLQY 225
DB 228 PKPIESIMTRCWSKDSQRPSPMEIYKIMTHLMRYFGADEPLQY 283

RESULT 8
Q9VCV0 PRELIMINARY; PRT; 252 AA.
ID Q9VCV0;
AC Q9VCV0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE CG4803 protein.
GN CG4803.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Eukaryota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxId=7227;
RX NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA ADAMS M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
```

ID	Q90ZY8	PRELIMINARY;	PRT;	371 AA.
AC	Q90ZY8;			
DT	01-DEC-2001 (TrEMBLrel. 19, Created)			
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)			
DT	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)			
DE	Protein kinase Npk.			
OS	Brachydanio rerio (Zebrafish) (Zebra danio).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;			
OC	Cyprinidae; Danio.			
OX	NCBI_TaxID=7955;			
RP	[1]			
RN	SEQUENCE FROM N.A.			
RC	Chou C.-M., Lee I.-L., Leu J.-H., Huang C.-J.;			
RT	"A novel protein kinase, zNPK, from the zebrafish.";			
RL	Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: AP265343; GAK52416.1; -			
DR	InterPro: IPR000719; Euk_pkinase.			
DR	InterPro: IPR002290; Ser_thr_pkinase.			
DR	Pfam: PF00069; pkinase; 1.			
DR	ProDom: PD00001; Euk_pkinase; 1.			
DR	PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.			
DR	PROSITE: PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.			
KW	ATP-binding; Kinase; Transferase.			
SQ	SEQUENCE 371 AA; 42456 MW; 9B918B8AB20D296 CRC64;			
	Query Match 30.0%; Score 376; DB 13; Length 371;			
	Best Local Similarity 39.0%; Pred. No. 7.4e-30;			
	Matches 82; Conservative 37; Mismatches 81; Indels 10; Gaps			
Qy	2 ELRQLSRVNHPTVKLYGACINPV--CLVMEYAEGGSLYNLHGAEPLPYTTAAHMSWC 59			
Db	84 EAEILSVLSHKNIQFYGAILEAPNDGIVTEYASRGSLEYLSSADS-EEEMDQVMTWA 142			
Qy	60 LQSQGVAYLHSHQPKALIHRLDKPNNLLVAGTVLKICDFGTACDIQTHMTNKK--GS 117			
Db	143 METAAGMHHYLHAEAPLVKVIHRLDKSRNVLTA-DNVLKICDFG-ASKMVSHTTHMSLVGT 200			
Qy	118 AAMWAPVFGSVSEKCDVFSWGIIWEIVTRRKPFDEIGGPAFRIMW-AVHNGTRPPL 176			
Db	201 FPMWAPVIGSLPVSETCDIYSVGVVLWEMLTREVFPK--GFGLQVAVLWVVEKHERPTI 258			
Qy	177 IKNLKPKIESLMTRCWSKPSQSRPSEIIV 206			
Db	259 PSSCPASFADLMRRCWNAEPKRPQPKQIL 288			
RESULT 10				
Q9ZQ31	Q9ZQ31	PRELIMINARY;	PRT;	411 AA.
ID	Q9ZQ31			
AC	Q9ZQ31; Q94A10.			
DT	01-MAY-1999 (TrEMBLrel. 10, Created)			
DT	01-JUN-2002 (TrEMBLrel. 21, Last sequence update)			
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)			
DE	Hypothetical 46.0 kDa protein.			
GN	AT2G24360.			
OS	Arabidopsis thaliana (Mouse-ear cress).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsi.			
OX	NCBI_TaxID=3702;			
RN	[1]			
RN	SEQUENCE FROM N.A.			
RC	STRAIN=CV. COLUMBIA;			
RA	Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanAken S.E.,			
RA	Barnstead M.E., Mason T.M., Bowman C.L., Ronning C.M., Benito M.-I.,			
RA	Carraer A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,			
RA	Fraser C.M., Venter J.C.;			
RL	Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.			
RP	[2]			
RN	SEQUENCE FROM N.A.			
RA	Yamada K., Liu S.X., Pham P.K., Banh J., Dale J.M., Gibson H.A.,			
RA	Goldsmith A.D., Jiang P.X., Lee J.M., Onodera C.S., Ouach H.L.,			

RA Tang C., Toriumi M., Yu G., Yu S., Bowser L., Carinici P., Chen H.,
RA Cheuk R., Hayashizaki Y., Ichida J., Jones T., Kamya A., Lin J.,
RA Karlin-Neumann G., Kawai J., Kim C., Koesema E., Lam B., Lin J.,
RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.D.,
RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Tracy S.E.,
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.,
RT "Full length cDNA of gene trs124.9/Atc2g24360 (GI:4337195).";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC006403; M18109.2; -.
DR EMBL: AY046026; AAK76700.1; -.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; UNKNOWN; 1.
DR Hypochemical protein; ATP-binding; Transferase.
SQ SEQUENCE 411 AA; 46001 MW; 3B7001CFBB411BAA CRC64;

Query Match 29.9%; Score 374; DB 10; Length 411;
Best Local Similarity 39.4%; Pred. No. 1.3e-29;
Matches 86; Conservative 40; Mismatches 70; Indels 22; Gaps 9;

QY 2 ELRLSRVNHPIVLYGACLPV--CLVMEYAGGSLYVNL---HGAEPLPYTAHA 55
DB 179 EVSMLANLKHPIVIRFGACIKPMWCIVTEYAKGGSVRFTRQRNAVPLKL-----A 233
QY 56 MSWCLQSGVAYVYHSMQPKALIHRLDKPNNLLVAGGTVLKICDFTG-CDIQTH-MTN 113
DB 234 VKQALDVARGAAYVHG---RNFIHRDLKSDNLLISADKSI-KIADFGVARIIEVQTEGMTP 289
QY 114 NKSSAAMAEVEFGESNYSEKCVFSWGIIIMVITRRKPFDEIGG--PAFRIMAVNHG 171
DB 290 ETGYRMWAEPMIOHRAPIYQKVDVYSGIVLWELITGLPQNMVAVQAFAV---VNRG 346
QY 172 TRPPLIKNLPKPIESLMTRCMSKDPSPQPSMEIYKIM 209
DB 347 VRPTVPADCLPVLGEMITRCMDADPEVRPCFAEIVNL 384

RESULT 11
ID Q9M085 PRELIMINARY; PRT; 412 AA.
AC Q9M085;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Protein kinase-like protein.
GN AT4G31170
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroside II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Leonard N., Quail M., Harris B., Rajandream M.A., Barrell B.G.,
RA Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: AL161578; CAB79835.1; -.
DR HSP: P08631; IAD5.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR004040; STY_pkinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Euk_pkinase; 1.

DR SMART: SM00221; STYKc; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 412 AA; 46083 MW; 644F35A90210D48 CRC64;

Query Match 29.8%; Score 373; DB 10; Length 412;
Best Local Similarity 39.9%; Pred. No. 1.7e-29;
Matches 87; Conservative 39; Mismatches 70; Indels 22; Gaps 9;

QY 2 ELRLSRVNHPIVLYGACLPV--CLVMEYAGGSLYVNL---HGAEPLPYTAHA 55
DB 180 EVSMLANLKHPIVIRFGACIKPMWCIVTEYAKGGSVRFTRQRNAVPLKL-----A 234
QY 56 MSWCLQSGVAYVYHSMQPKALIHRLDKPNNLLVAGGTVLKICDFTG-CDIQTH-MTN 113
DB 235 VMQALDVARGAAYVHG---RNFIHRDLKSDNLLISADKSI-KIADFGVARIIEVQTEGMTP 290
QY 114 NKSSAAMAEVEFGESNYSEKCVFSWGIIIMVITRRKPFDEIGG--PAFRIMAVNHG 171
DB 291 ETGYRMWAEPMIOHRAPIYQKVDVYSGIVLWELITGLPQNMVAVQAFAV---VNRG 347
QY 172 TRPPLIKNLPKPIESLMTRCMSKDPSPQPSMEIYKIM 209
DB 348 VRPTVPADCLPVLGEMITRCMDADPEVRPCFAEIVNL 385

RESULT 12
ID Q9ESL3 PRELIMINARY; PRT; 454 AA.
AC Q9ESL3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE MLTK-beta.
GN ZAK OR MLTK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21264927; Pubmed=11042189;
RA Gotoh I., Adachi M., Nishida E.;
RT "Identification and Characterization of a Novel MAP Kinase Kinase
RT Kinase, MLTK.";
RL J. Biol. Chem. 276:4276-4286(2001).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: AB049732; BAB16443.1; -.
DR HSP: P12931; 1PMK.
DR MGD: MGI:1931274; Zak.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR004040; STY_pkinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00221; STYKc; 1.
DR SMART: SM00220; S_TKc; 1.
DR SMART: SM00219; TYTKc; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR ATP-binding; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 454 AA; 51366 MW; 35C2FCD0729D9395 CRC64;

Query Match 29.7%; Score 372; DB 11; Length 454;
Best Local Similarity 37.7%; Pred. No. 2.4e-29;
Matches 81; Conservative 40; Mismatches 80; Indels 14; Gaps 8;

QY 2 ELRLSRVNHPIVLYGACLPV--CLVMEYAGGSLYVNLG--AEPLPYTAHA 57
DB 53 EALILSVLSHRNITQFYGVILEPPIYGVITEYASLSGLYINSNSEE--DMEHIMT 109

Qy	58	WCLQCSQGVAYLHSMQPKALIHRLDKPNNLLLVAGGTVLVKICDFGTACDIQTHMTNNK--	111
Db	110	WATDVAKGMYLHMEAPVKVIHRDLKSRNVVIAADG-VLKICDFG-ASRPHNHTTHMSLV	167
Qy	116	GSAAWMAPEVFEGSNTSEKCDVFSWGIIWEVITRRKPFDEIGGPAFRIMW-AVHNGTRP	174
Db	168	GTFPWMAPEVIQSLPVSSETCDTYSYGVVLWEMLTREVPFKGLEG--LQVAVLWVEKNERL	225
Qy	175	PLIKNLPKPIESLMTRCWSKDPQRSMEIEIVKIM	209
Db	226	TIPSSCSPFAELLHQCEADAKRPSFKQIISIL	260
RESULT 13			
Q9ESL4			
ID	Q9ESL4	PRELIMINARY; PRT; 802 AA.	
AC	Q9ESL4;		
DT	01-MAR-2001	(T-EMBLrel. 16, Created)	
DT	01-MAR-2001	(T-EMBLrel. 16, Last sequence update)	
DT	01-JUN-2002	(T-EMBLrel. 21, Last annotation update)	
DE	MUTK alpha.		
GN	ZAK OR MLTK.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RP	MEDLINE=21264927; PubMed=11042189;		
RT	Gotoh I., Adachi M., Nishida E.;		
RT	"Identification and Characterization of a Novel MAP Kinase Kinase		
RT	Kinase, MLTK.",		
RT	J. Biol. Chem. 276:4276-4286(2001).		
CC	-1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.		
DR	EMBL; AB049731; BAB16442.1; -.		
DR	HSSP; P12931; 1PMK.		
DR	MGD; MGI:1931274; Zak.		
DR	InterPro; IPR000194; ATPase_a/bcentre.		
DR	InterPro; IPR000719; Euk_pkinase.		
DR	InterPro; IPR001660; SAM.		
DR	InterPro; IPR002290; Ser_thr_pkinase.		
DR	InterPro; IPR004040; STY_pkinase.		
DR	Pfam; PF00069; pkinase; 1.		
DR	Pfam; PF00336; SAM; 1.		
DR	PRINTS; PR00109; TYRKINASE.		
DR	ProDom; PD000001; Euk_pkinase; 1.		
DR	SMART; SM00454; SAM; 1.		
DR	SMART; SM00221; STYKC; 1.		
DR	SMART; SM00220; S_TKC; 1.		
DR	SMART; SM00219; TYRK; 1.		
DR	PROSITE; PS00152; ATPASE_ALPHA_BETA; UNKNOWN 1.		
DR	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.		
DR	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.		
KW	ATP-binding; Serine/threonine-protein kinase; Transferase.		
SEQ	SEQUENCE 802 AA; 91719 MW; D431DPF8F312A43CC CRC64;		
Query Match 29.7%; Score 372; DB 11; Length 802;			
Best Local Similarity 37.7%; Pred. No. 4.8e-29;			
Matches 81; Conservative 40; Mismatches 80; Indels 14; Gaps			
Qy	2	ELRLQSRVNHNPVTKVLYGACLNIP--VCLWMEYAEAGSLYNVLHG--AEPLPYTTAAHAMS	57
Db	53	EAILSVLSHRNLIQFYGVILPEPNVGI VTEYASLGSGLYDINSKSEEM---DMEHMT	109
Qy	58	WCLQCSQGVAYLHSMQPKALIHRLDKPNNLLLVAGGTVLVKICDFGTACDIQTHMTNNK--	115
Db	110	WATDVAKGMYLHMEAPVKVIHRDLKSRNVVIAADG-VLKICDFG-ASRPHNHTTHMSLV	167
Qy	116	GSAAWMAPEVFEGSNTSEKCDVFSWGIIWEVITRRKPFDEIGGPAFRIMW-AVHNGTRP	174
Db	168	GTFPWMAPEVIQSLPVSSETCDTYSYGVVLWEMLTREVPFKGLEG--LQVAVLWVEKNERL	225

```

Qy  58  MCLQSGCAVVAVYHSNQFPALHRLDLKPNPLLVLGVGLFKICDPTGACDIIQTHNTNKK--  115
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  110  WATDVAKGHHYVHEMAEPKAVVHLHRDKSKNVVIAADG-VLKI.CDYG-ASRFNNHTTHSLV  167
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy  116  GSAAMNAPEVEFGGSNYSEKCVFSWGILMEVITRRKPFDEIGGAPAFIMW-AVHNQTRP  174
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  168  GTFPMMAREVIOLEPVSCTDTIYSGVVLNEMLTREVPFKGLEG--LQVAMLVVEKNERL  225
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy  175  PLIKULPKPIESIMTRCWSKDPSCQRFSEMEIIVKM  209
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  226  TTPSSCPRSFAELHQCMEADAKKRPSPKQIISLI  260

```

Search completed: December 10, 2002, 03:58:25
Job time : 33 secs

RESULT	15
Q9HDD2	
ID	Q9HDD2
AC	Q9HDD2;
DT	01-MAR-2001 (TrEMBLrel. 16, Created)
DT	01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE	Placable mixed-lineage kinase protein.
GN	MLKLAK.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homn.
OX	NCBI_TaxID=9606;
OX	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=LYMPHOID ORGAN;
RA	Abe Y., Ueda N.;
RT	"Placible mixed-lineage kinase derived from LAK cell.";
RL	Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC	!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR	EMBL; AB030034; BAB12040.1; -.
DR	HSSP; P12931; IFMK.
DR	InterPro: IPR000719; Euk.pkinase.
DR	InterPro: IPR001660; SAM.
DR	InterPro: IPR002290; Ser_thr.pkinase.
DR	InterPro: IPR004040; STY_pkinase.
DR	InterPro: IPR001245; TYR_pkinase.
DR	Pfam; PF00069; pkinase; 1.
DR	Pfam; PF00536; SAM; 1.
DR	PRINTS; PRO0109; TYRKINASE.
DR	ProDom; PD000001; Euk.pkinase; 1.
DR	SMART; SM00454; SAM; 1.
DR	SMART; SM00221; STYK; 1.
DR	SMART; SM00220; S_TKC; 1.
DR	SMART; SM00219; TYRK; 1.
DR	PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
DR	PROSITE; PSS00108; PROTEIN_KINASE_ST; 1.
DR	ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase
SQ	SEQUENCE 800 AA; 9155 MW; B2814509EC54B07A CRC64;

Query Match	29.6%;	Score 371;	DB 4;	Length 800;
Best Local Similarity	37.7%;	Pred. No. 6e-29;		
Matches	81;	Conservative	40;	Mismatches 80;
				Indels 14;
				Gaps 8;

[illegible]

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 10, 2002, 03:56:30 ; Search time 19 Seconds
(without alignments)
353.075 Million cell updates/sec

Title: US-09-830-144-2_COPY_76_303

Perfect score: 1252
Sequence: 1 VELRQLSRVNHPIVLYGA.....MTHLMRYPPGADEPLQVPCQ 228

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/1/iaa/5A_COMB.pep:*
- 2: /cgn2_6/prodata/1/iaa/5B_COMB.pep:*
- 3: /cgn2_6/prodata/1/iaa/6A_COMB.pep:*
- 4: /cgn2_6/prodata/1/iaa/6B_COMB.pep:*
- 5: /cgn2_6/prodata/1/iaa/PCTUS_COMB.pep:*
- 6: /cgn2_6/prodata/1/iaa/backfiles.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1252	100.0	579	US-09-529-279-4	Sequence 4, Appli
2	1252	100.0	590	US-09-529-279-15	Sequence 15, Appli
3	371	29.6	455	US-09-221-235-5	Sequence 5, Appli
4	371	29.6	455	US-09-221-928-5	Sequence 5, Appli
5	371	29.6	455	US-09-221-527-5	Sequence 5, Appli
6	371	29.6	455	US-09-221-236-5	Sequence 5, Appli
7	371	29.6	455	US-09-221-416-5	Sequence 5, Appli
8	371	29.6	455	US-09-221-245-5	Sequence 5, Appli
9	371	29.6	455	US-09-163-115-5	Sequence 5, Appli
10	371	29.6	455	US-09-221-528-5	Sequence 5, Appli
11	371	29.6	455	US-09-593-553-5	Sequence 5, Appli
12	371	29.6	455	US-09-221-237-5	Sequence 5, Appli
13	353.5	28.2	668	US-08-205-018-2	Sequence 2, Appli
14	353.5	28.2	859	US-08-395-580-2	Sequence 2, Appli
15	353.5	28.2	859	PCT-US95-02792-2	Sequence 2, Appli
16	338.5	27.0	835	US-09-291-839-2	Sequence 2, Appli
17	323	25.8	821	US-07-928-464-2	Sequence 2, Appli
18	323	25.8	821	US-08-003-311B-2	Sequence 2, Appli
19	323	25.8	821	US-08-261-432-2	Sequence 2, Appli
20	323	25.8	821	PCT-US93-07347-2	Sequence 2, Appli
21	320	25.6	263	US-09-035-706-5	Sequence 5, Appli
22	320	25.6	263	US-08-955-841-5	Sequence 5, Appli
23	320	25.6	263	US-09-390-425-5	Sequence 5, Appli
24	320	25.6	263	US-09-566-906-5	Sequence 5, Appli
25	309.5	24.7	269	US-07-857-224B-79	Sequence 79, Appli
26	307.5	24.6	276	US-07-857-224B-72	Sequence 72, Appli
27	306	24.4	275	US-07-857-224B-71	Sequence 71, Appli

28	305.5	24.4	787	4	US-09-188-930-334	Sequence 334, App
29	304.5	24.3	304	2	US-08-701-191A-27	Sequence 27, Appl
30	302.5	24.2	316	1	US-08-278-089A-16	Sequence 16, Appl
31	302.5	24.2	316	2	US-08-838-957A-15	Sequence 15, Appl
32	299.5	23.9	527	4	US-08-426-509A-10	Sequence 10, Appl
33	299.5	23.9	527	5	PCT-US95-05008-10	Sequence 10, Appl
34	299.5	23.9	625	1	US-08-391-615-3	Sequence 3, Appli
35	298.5	23.8	620	4	US-08-426-509A-9	Sequence 9, Appli
36	298.5	23.8	620	5	PCT-US95-05008-9	Sequence 9, Appli
37	294.5	23.5	293	2	US-08-701-191A-26	Sequence 26, Appl
38	294	23.5	536	4	US-09-188-930-185	Sequence 185, App
39	293.5	23.4	274	2	US-08-469-537A-55	Sequence 55, Appl
40	290	23.2	1584	4	US-09-457-040B-27	Sequence 27, Appl
41	284.5	22.7	293	2	US-08-701-191A-16	Sequence 16, Appl
42	280	22.4	388	1	US-08-070-165F-4	Sequence 4, Appli
43	280	22.4	388	2	US-08-885-418-4	Sequence 5, Appli
44	278.5	22.2	588	1	US-08-391-615-5	Sequence 5, Appli
45	278	22.2	299	2	US-08-701-191A-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1

US-09-529-279-4
; Sequence 4, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529, 279
; CURRENT FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-279-4

Query Match	100.0%;	Score 1252;	DB 4;	Length 579;
Best Local Similarity	100.0%;	Pred. No. 3.2e-133;		
Matches 228;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	VELRQLSRVNHPIVLYGACLNPCVLVMEYAEAGGSLYVNLHGAEPLPYVYTAAHAMSWCL	60	
Db	76	VELRQLSRVNHPIVLYGACLNPCVLVMEYAEAGGSLYVNLHGAEPLPYVYTAAHAMSWCL	135	
Qy	61	QCSQGVAYLHSMQPKALIHRLDKPPNLLLVAGGTVLXICDFGTACDIQTHMTNNKGSAAW	120	
Db	136	QCSQGVAYLHSMQPKALIHRLDKPPNLLLVAGGTVLXICDFGTACDIQTHMTNNKGSAAW	195	
Qy	121	MAPEVFGSNYSKCDVFSNGIILWEVITRRKPPDEIGGPAFRIMMAVHNGTRPPLIKNL	180	
Db	196	MAPEVFGSNYSKCDVFSNGIILWEVITRRKPPDEIGGPAFRIMMAVHNGTRPPLIKNL	255	
Qy	181	PKPIESLMTRCWSKDPORPSMEEIVKIMTHLMRYFFGADEPLQVPCQ	228	
Db	256	PKPIESLMTRCWSKDPORPSMEEIVKIMTHLMRYFFGADEPLQVPCQ	303	
RESULT 2				
US-09-529-279-15				
; Sequence 15, Application US/09529279				
; Patent No. 6451617				
; GENERAL INFORMATION:				

APPLICANT: ONO, KOICHIRO
APPLICANT: OHTOMO, TOSHIHIKO
APPLICANT: TSUCHIYA, MASAYUKI
TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
FILE REFERENCE: 053466/0278
CURRENT APPLICATION NUMBER: US/09/529,279
CURRENT FILING DATE: 2000-04-11
PRIOR APPLICATION NUMBER: PCT/JP98/04796
PRIOR FILING DATE: 1998-10-22
PRIOR APPLICATION NUMBER: JP 9/290188
PRIOR FILING DATE: 1997-10-22
NUMBER OF SEQ ID NOS: 48
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 15
LENGTH: 590
TYPE: PRT
ORGANISM: Homo sapiens
US-09-529-279-15

Query Match 100.0%; Score 1252; DB 4; Length 590;
Best Local Similarity 100.0%; Pred. No. 3,3e-133;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VELRLSRVNHPIVLYGACLNPCVCLNMEYAEAGSLYVNLHGAEPLPYTTAAHMSWCL 60
DB 76 VELRLSRVNHPIVLYGACLNPCVCLNMEYAEAGSLYVNLHGAEPLPYTTAAHMSWCL 135
QY 61 QCSQGVAVYHSMQKALIHRLDKPNNLLVAGGTVLKICDFGTACDIQTHMTNKGSAAM 120
DB 136 QCSQGVAVYHSMQKALIHRLDKPNNLLVAGGTVLKICDFGTACDIQTHMTNKGSAAM 195
QY 121 MAEVEGSGSYSEKCVFSGIILMEVITRRKPFDEIGAPRIMAWHNGTRPPLIKL 180
DB 196 MAEVEGSGSYSEKCVFSGIILMEVITRRKPFDEIGAPRIMAWHNGTRPPLIKL 255
QY 181 PKRISLMTRCWSKDPSPRSMEEIVKIMTHLMRYFPGADEPLQYPCQ 228
DB 256 PKRISLMTRCWSKDPSPRSMEEIVKIMTHLMRYFPGADEPLQYPCQ 303

RESULT 3
US-09-221-235-5
Sequence 5, Application US/09221235
Patent No. 6043040
GENERAL INFORMATION:
APPLICANT: Acton, Susan
TITLE OF INVENTION: NOVEL CSAK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
FILE REFERENCE: NMI-050
CURRENT APPLICATION NUMBER: US/09/221,235
CURRENT FILING DATE: 1998-12-28
EARLIER APPLICATION NUMBER: 09/163,115
EARLIER FILING DATE:
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 455
TYPE: PRT
ORGANISM: Homo sapiens
US-09-221-235-5

Query Match 29.6%; Score 371; DB 3; Length 455;
Best Local Similarity 37.7%; Pred. No. 1,4e-33;
Matches 81; Conservative 40; Mismatches 80; Indels 14; Gaps 8;

QY 2 ELRLSRVNHPIVLYGACLNPCVCLNMEYAEAGSLYVNLHGAEPLPYTTAAHMS 57
DB 53 EALILSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGLYDINSRSEEM---DMDHIMT 109
QY 58 WCLQCSQGVAVYHSMQKALIHRLDKPNNLLVAGGTVLKICDFGTACDIQTHMTNKG 115
DB 110 WATDVAKGMHLYHAEAVKVIHRLDKSRNVVIAADG-VLKICDFG-ASRFHNTTHMSLV 167
QY 116 GSAAMMAPEVEGSGSYSEKCVFSGIILMEVITRRKPFDEIGAPRIMAWHNGTRP 174

DB 168 GTFPMMAPEVIGSLPVSECTDITYSYGVLMEMLTREVPFGLEG--LQVAMLYVEKNERL 225
QY 175 PLINLKPPIESLMTRCWSKDPSPRSMEEIVKIM 209
DB 226 TIPSCPRFAELHLQCEWADAKKRPFSKQIISIL 260

RESULT 4
US-09-221-928-5
Sequence 5, Application US/09221928
Patent No. 6121030
GENERAL INFORMATION:
APPLICANT: Acton, Susan
TITLE OF INVENTION: NOVEL CSAK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
FILE REFERENCE: NMI-050
CURRENT APPLICATION NUMBER: US/09/221,928
CURRENT FILING DATE: 1998-12-28
EARLIER APPLICATION NUMBER: 09/163,115
EARLIER FILING DATE:
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 455
TYPE: PRT
ORGANISM: Homo sapiens
US-09-221-928-5

Query Match 29.6%; Score 371; DB 3; Length 455;
Best Local Similarity 37.7%; Pred. No. 1,4e-33;
Matches 81; Conservative 40; Mismatches 80; Indels 14; Gaps 8;

QY 2 ELRLSRVNHPIVLYGACLNPCVCLNMEYAEAGSLYVNLHGAEPLPYTTAAHMS 57
DB 53 EALILSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGLYDINSRSEEM---DMDHIMT 109
QY 58 WCLQCSQGVAVYHSMQKALIHRLDKPNNLLVAGGTVLKICDFGTACDIQTHMTNKG 115
DB 110 WATDVAKGMHLYHAEAVKVIHRLDKSRNVVIAADG-VLKICDFG-ASRFHNTTHMSLV 167
QY 116 GSAAMMAPEVEGSGSYSEKCVFSGIILMEVITRRKPFDEIGAPRIMAWHNGTRP 174
DB 168 GTFPMMAPEVIGSLPVSECTDITYSYGVLMEMLTREVPFGLEG--LQVAMLYVEKNERL 225
QY 175 PLINLKPPIESLMTRCWSKDPSPRSMEEIVKIM 209
DB 226 TIPSCPRFAELHLQCEWADAKKRPFSKQIISIL 260

RESULT 5
US-09-221-527-5
Sequence 5, Application US/09221527
Patent No. 6146832
GENERAL INFORMATION:
APPLICANT: Acton, Susan
TITLE OF INVENTION: NOVEL CSAK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
FILE REFERENCE: NMI-050
CURRENT APPLICATION NUMBER: US/09/221,527
CURRENT FILING DATE: 1998-12-28
EARLIER APPLICATION NUMBER: 09/163,115
EARLIER FILING DATE:
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 455
TYPE: PRT
ORGANISM: Homo sapiens
US-09-221-527-5

Query Match 29.6%; Score 371; DB 4; Length 455;
Best Local Similarity 37.7%; Pred. No. 1,4e-33;
Matches 81; Conservative 40; Mismatches 80; Indels 14; Gaps 8;

Qy	2	ELRQLSRVNHNPVVKLYGACLP--VCLMVEYABGSGSLVNVLHG--ASPLPYTTAAHAMS	57
Db	53	EAEILSVLSHRNIIQFYGVILEPPNNGIVTEYASGLSYDYINSRSEB---DMDHINT	109
Qy	58	WCLQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGTGLVKICDFGTACDIOIOTMTNNK--	115
Db	110	WATDVAKGWHYLHMEAPVKVHIRDLKSRNVVIAADG-VLKICDFG-ASRFNHHTHMSLV	167
Qy	116	GSAAMWAPVEFGSGNSYSEKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMW-AVHNGTRP	174
Db	168	GTFFPMWAEVIQSLPVSCTDIYSYGVVLWEMLTREVPFKGLEG--LQVAMLVVEKNERL	225
Qy	175	PLIKNLKPISLMTRCWSKDSQSPRSMBEIVKIM	209
Db	226	TIPSSCPRSFABLLHQCEADAKRPSFKQIISIL	260

RESULT 6
 US-09-221-236-5
 ; Sequence 5, Application US/09221236
 ; Patent No. 6146841
 ; GENERAL INFORMATION:
 ; APPLICANT: Acton, Susan
 ; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
 ; FILE REFERENCE: MNI-050
 ; CURRENT APPLICATION NUMBER: US/09/221.236
 ; CURRENT FILING DATE: 1998-12-28
 ; EARLIER APPLICATION NUMBER: 09/163,115
 ; EARLIER FILING DATE: 1998-09-29
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 5
 ; LENGTH: 455
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-221-236-5

Query Match 29.6%; Score 371; DB 4; Length 455;
 Best Local Similarity 37.7%; Pred. No. 1.4e-33;
 Matches 81; Conservative 40; Mismatches 80; Indels 14; Gaps 8;

Qy	2	ELRQLSRVNHNPVVKLYGACLP--VCLMVEYABGSGSLVNVLHG--ASPLPYTTAAHAMS	57
Db	53	EAEILSVLSHRNIIQFYGVILEPPNNGIVTEYASGLSYDYINSRSEB---DMDHINT	109
Qy	58	WCLQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGTGLVKICDFGTACDIOIOTMTNNK--	115
Db	110	WATDVAKGWHYLHMEAPVKVHIRDLKSRNVVIAADG-VLKICDFG-ASRFNHHTHMSLV	167
Qy	116	GSAAMWAPVEFGSGNSYSEKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMW-AVHNGTRP	174
Db	168	GTFFPMWAEVIQSLPVSCTDIYSYGVVLWEMLTREVPFKGLEG--LQVAMLVVEKNERL	225
Qy	175	PLIKNLKPISLMTRCWSKDSQSPRSMBEIVKIM	209
Db	226	TIPSSCPRSFABLLHQCEADAKRPSFKQIISIL	260

RESULT 7
 US-09-221-416-5
 ; Sequence 5, Application US/09221416
 ; Patent No. 6153417
 ; GENERAL INFORMATION:
 ; APPLICANT: Acton, Susan
 ; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
 ; FILE REFERENCE: MNI-050
 ; CURRENT APPLICATION NUMBER: US/09/221.416
 ; CURRENT FILING DATE: 1998-12-28
 ; EARLIER APPLICATION NUMBER: 09/163,115
 ; EARLIER FILING DATE: 1998-09-29
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 5


```

; TITLE OF INVENTION: Sequences Encoding the Same and Methods Related Thereto
;
; NUMBER OF SEQUENCES: 4
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and Norris
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch disk, 720 KB
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect 5.1
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02792
; FILING DATE: herewith
; CLASSIFICATION:
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/205,018
; FILING DATE: 01-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Rebecca L. Ralph (formerly Gaumond)
; REGISTRATION NUMBER: 35,152
; REFERENCE/DOCKET NUMBER: CH-0488
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
;
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 859 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-02792-2

```

Query Match 28.2%; Score 353.5; DB 5; Length 859;
Best Local Similarity 34.7%; Pred. No. 3.2e-31;
Matches 75; Conservative 43; Mismatches 81; Indels 17; Gaps 8;

```

QY 2 ELRLSRVNHPIVVKLYGACLNLP--VCLVMEYAEAGSLYVNLHGAPLPYPTAAHAMSWC 59
DB 161 DIKHLKRLKHPNITTFKGYCTQAPCYCIMEFCAQGLYEVLRAGRPF--TPSLLVDS 217
QY 60 LQCSQGVAVYLSHMQPKALIHRLDKPPNLLIVAGTVLKI CDFTACDIQTHMTNKK--GS 117
DB 218 MGJAGGMNVLHLHK--IIHRDLKSPN-MLITYDVVKISDFGTSKELSDKSTRKSFAGT 273
QY 118 AAMMAPEVFEFGSNYSKCDVFSWGIILWEVITRRKPDDEIGPAPRIMAV--HNGTRPPL 176
DB 274 VAMMAPEVIRNEPVSEKVDIWSFGVVLWELTGEIPIKVDSSA--IIMGVGSNSLHLPV 331
QY 177 IKNLKPKIESLMTKWSKDPQSPRSMEEIVKIMTHL 212
DB 332 PSSCPDGFKILLRQCNWSKRPNRPSFRQ--ILLHL 364

```

Search completed: December 10, 2002, 03:59:30
Job time : 21 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 10, 2002, 03:57:20 ; Search time 10 Seconds
(without alignments)
370.325 Million cell updates/sec

Title: US-09-830-144-2_COPY_76_303
Perfect score: 1252
Sequence: 1 VELRQLSRVNHNPVIVKLYGA.....MTHLMRYFFCADEPLOYPCQ 228

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 103943 seqs, 16242309 residues
Total number of hits satisfying chosen parameters: 103943

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pbp:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pbp:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pbp:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pbp:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pbp:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pbp:*
7: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pbp:*
8: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pbp:*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pbp:*
10: /cgn2_6/ptodata/1/pubpaa/US05_PUBCOMB.pbp:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pbp:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pbp:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pbp:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pbp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1252	100.0	579	9	US-10-158-895-4
2	1252	100.0	590	9	US-10-158-895-15
3	371	29.6	455	10	US-09-757-982-5
4	363.5	29.0	394	10	US-09-862-027-19
5	355	28.4	1036	12	US-10-014-882-2
6	346	27.6	328	10	US-09-862-027-18
7	338.5	27.0	835	10	US-09-947-199-2
8	335.5	26.8	835	10	US-09-947-199-8
9	331.5	26.5	966	10	US-09-771-161A-197
10	320	25.6	263	10	US-09-840-704-5
11	313.5	25.0	850	10	US-09-904-389-2
12	312	24.9	265	10	US-09-797-039-10
13	312	24.9	277	10	US-09-815-915-13
14	312	24.9	277	10	US-09-882-166-4
15	312	24.9	278	9	US-09-842-582-4
16	312	24.9	278	10	US-09-797-039-13
17	312	24.9	278	10	US-09-922-138-18
18	312	24.9	278	10	US-09-922-138-27
19	312	24.9	278	10	US-09-910-150-17

20	312	24.9	278	10	US-09-910-150-31	Sequence 31, Appl
21	308	24.6	278	10	US-09-815-915-16	Sequence 16, Appl
22	307	24.5	237	10	US-09-916-790-17	Sequence 17, Appl
23	304	24.3	273	10	US-09-922-138-11	Sequence 11, Appl
24	303	24.2	247	10	US-09-916-790-7	Sequence 7, Appl
25	303	24.2	271	10	US-09-799-875-28	Sequence 28, Appl
26	303	24.2	272	10	US-09-780-949-4	Sequence 4, Appl
27	303	24.2	272	10	US-09-910-150-27	Sequence 27, Appl
28	301.5	24.1	279	10	US-09-799-875-23	Sequence 23, Appl
29	299.5	23.9	527	10	US-09-977-269-10	Sequence 10, Appl
30	298.5	23.8	620	10	US-09-977-269-9	Sequence 9, Appl
31	296	23.6	272	10	US-09-815-915-14	Sequence 14, Appl
32	292.5	23.4	280	10	US-09-515-806-15	Sequence 15, Appl
33	289	23.1	425	10	US-09-828-313-29	Sequence 29, Appl
34	288.5	23.0	628	10	US-09-862-027-48	Sequence 48, Appl
35	282.5	22.6	627	10	US-09-862-027-47	Sequence 47, Appl
36	278	22.2	310	9	US-09-939-833-7	Sequence 7, Appl
37	278	22.2	310	10	US-09-939-754-7	Sequence 7, Appl
38	278	22.2	310	10	US-09-939-832-7	Sequence 7, Appl
39	278	22.2	764	10	US-09-925-302-714	Sequence 714, App
40	273.5	21.8	214	10	US-09-515-806-10	Sequence 10, Appl
41	271.5	21.7	239	10	US-09-797-039-12	Sequence 12, Appl
42	271.5	21.7	245	10	US-09-815-915-15	Sequence 15, Appl
43	270.5	21.6	212	10	US-09-834-496A-4	Sequence 4, Appl
44	270.5	21.6	214	10	US-09-860-352A-6	Sequence 6, Appl
45	269.5	21.5	507	10	US-09-977-269-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-10-158-895-4
; Sequence 4, Application US/10158895
; Patent No. US20020155624A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: TSUCHIYA, MASAYUKI
; APPLICANT: OTOMO, TOSHIKO
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10158,895
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-158-895-4

Query Match 100.0%; Score 1252; DB 9; Length 579;
Best Local Similarity 100.0%; Pred. No. 2.7e-104;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	VELRQLSRVNHNPVIVKLYGACLNVPVCLVMEYAEAGGSLYNVLHGAEPLVPYTTAAHMSWCL	60
Db	76	VELRQLSRVNHNPVIVKLYGACLNVPVCLVMEYAEAGGSLYNVLHGAEPLVPYTTAAHMSWCL	135
Qy	61	QCSQGVAYLHNSQPKALIHRLDKPPNLLIVAGGTVLKICDFGTACDIQTHMTNNKGSAAW	120
Db	136	QCSQGVAYLHNSQPKALIHRLDKPPNLLIVAGGTVLKICDFGTACDIQTHMTNNKGSAAW	195
Qy	121	MAPEVFEAGNSYSEKCDVFSWGIIILMEVITRRKPPFDEIGGPAFRIMWAVHNGTRPPLIKNL	180
Db	196	MAPEVFEAGNSYSEKCDVFSWGIIILMEVITRRKPPFDEIGGPAFRIMWAVHNGTRPPLIKNL	255

Qy 181 PKPIESLMTRCWSKDPSPQSPMEIYKIMTHLMRYFPAGDEPLQYPCQ 228
|||
Db 256 PKPIESLMTRCWSKDPSPQSPMEIYKIMTHLMRYFPAGDEPLQYPCQ 303

RESULT 2

US-10-158-895-15
; Sequence 15, Application US/10158895
; Patent No. US20020158624A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 590
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-158-895-15

Query Match 100.0%; Score 1252; DB 9; Length 590;
Best Local Similarity 100.0%; Pred. No. 2.8e-104;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VELRLSRVNHPIVLYGACLNPCVLMEYAEAGSLYVNLHGAEPPLPYTTAAHMSWCL 60
|||
Db 76 VELRLSRVNHPIVLYGACLNPCVLMEYAEAGSLYVNLHGAEPPLPYTTAAHMSWCL 135
|||
Qy 61 QCSQGVAYLHSMQPKALIHRLDKPNLLVAGGTVLKICDFTACDIQTHMTNNKGSAAW 120
|||
Db 136 QCSQGVAYLHSMQPKALIHRLDKPNLLVAGGTVLKICDFTACDIQTHMTNNKGSAAW 195
|||
Qy 121 MADEVFEGSNYSKCDVFSWGIIIMVEVITRRKPFDEIGGAPFRIMAVHNGTRPPLIKL 180
|||
Db 196 MADEVFEGSNYSKCDVFSWGIIIMVEVITRRKPFDEIGGAPFRIMAVHNGTRPPLIKL 255
|||
Qy 181 PKPIESLMTRCWSKDPSPQSPMEIYKIMTHLMRYFPAGDEPLQYPCQ 228
|||
Db 256 PKPIESLMTRCWSKDPSPQSPMEIYKIMTHLMRYFPAGDEPLQYPCQ 303
|||

RESULT 3

US-09-757-982-5
; Sequence 5, Application US/09757982
; Patent No. US20020094559A1
; GENERAL INFORMATION:
; APPLICANT: Accon, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/757,982
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 09/163,115
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-757-982-5

Query Match 29.6%; Score 371; DB 10; Length 455;

Best Local Similarity 37.7%; Pred. No. 7.4e-26;
Matches 81; Conservative 40; Mismatches 80; Indels 14; Gaps 8;

Qy 2 ELRLSRVNHPIVLYGACLNPCVLMEYAEAGSLYVNLHGAEPPLPYTTAAHMS 57
|||
Db 53 EAEILSVLSHRNIIFQYGVILEPPNYGIVTEYASLGLDYINSRSEEM--DMDHIMT 109
|||
Qy 58 WCLQSGVAYLHSMQPKALIHRLDKPNLLVAGGTVLKICDFTACDIQTHMTNNK-- 115
|||
Db 110 WATDVAKGMHYLHMEAEVKVILHRLDKSRNVIAADG-VLKI CDFG-ASRPHNHTTMSLV 167
|||
Qy 116 GSAAMMAPEVEFGSNYSKCDVFSWGIIIMVEVITRRKPFDEIGGAPFRIMAVHNGTRP 174
|||
Db 168 GTFPMMAPEVIQSPVSETCDTYSYGVLMEMLTREVPFGLGEG--LQVAMLVENRNL 225
|||
Qy 175 PLIKNLPKPIESLMTRCWSKDPSPQSPMEIYKIM 209
|||
Db 226 TIPSQCRSPFAELHQCWEADAKKRPSPFQIISIL 260
|||

RESULT 4

US-09-862-027-19
; Sequence 19, Application US/09862027
; Patent No. US20020142428A1
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin R.
; TITLE OF INVENTION: No. US20020142428A1 Kinases and Uses Thereof
; FILE REFERENCE: 35800/234862
; CURRENT APPLICATION NUMBER: US/09/862,027
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 09/345,473
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-862-027-19

Query Match 29.0%; Score 363.5; DB 10; Length 394;
Best Local Similarity 37.7%; Pred. No. 2.9e-25;
Matches 83; Conservative 36; Mismatches 84; Indels 17; Gaps 6;

Qy 2 ELRLSRVNHPIVLYGACLNPCVLMEYAEAGSLYVNLHGAEPPLPYTTAAHMSWC 59
|||
Db 50 EAKLFAMLKHPNIIALRGVCLKEPNLCLVMEFARGGPLNRVLSGKXIPPDI---LVNMA 105
|||
Qy 60 LQCSQGVAYLHSMQPKALIHRLDKPNLLV-----AGTVLTKICDFTACD-IQTHM 111
|||
Db 106 VQIARGMNYLHDEAIVPIIHRDLKSNITLILQKVENGDLSNKKILKIDTDFELAREMHRITK 165
|||
Qy 112 TNNKGSAAAMMAPEVEFGSNYSKCDVFSWGIIIMVEVITRRKPFDEIGGAPFRIMAV-HN 170
|||
Db 166 MSAAGTYAMMAPEVIRASMSKSDVSYGVLLMELLTGEVPRRGIDG--LRVAYGVAMN 223
|||
Qy 171 GTRPPLIKNLPKPIESLMTRCWSKDPSPQSPMEIYKIMT 210
|||
Db 224 KLAIPISTCEPFAKLMEDCWNPDPSRPSFTIILDLT 263
|||

RESULT 5

US-10-014-882-2
; Sequence 2, Application US/10014882
; Patent No. US20020107384A1
; GENERAL INFORMATION:
; APPLICANT: Hu, Yi
; APPLICANT: Kleke, James
; APPLICANT: Donoho, Gregory
; TITLE OF INVENTION: No. US20020107384A1 Human Kinase and Polynucleotides Encoding r
; FILE REFERENCE: LEX-0279-USA
; CURRENT APPLICATION NUMBER: US/10/014,882
; CURRENT FILING DATE: 2001-12-11


```
; PRIOR APPLICATION NUMBER: US 60/254,744
; PRIOR FILING DATE: 2000-12-11
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1036
; TYPE: PRT
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1036)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-014-882-2

Query Match      28.4%; Score 355; DB 12; Length 1036;
Best Local Similarity 36.1%; Pred. No. 5e-24;
Matches 83; Conservative 40; Mismatches 85; Indels 22; Gaps 6;

QY 2 ELRLSRVNHPIVVKLYGACLN--VCLVMEYAEAGSLYNVLHGAEPLPYTTAAHMSW 52
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 171 EARLFAMLRHPIETELRGVCLQQPHLCVLLEFARGGALNALAAANAADPPRAGPRRAR 230
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 53 ---AHAM-SWCLOCSQGVAYLHNSQPKALIHRLDKPNNLLVAG-----GTVLKICDF 101
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 231 RIPPHVLNVAVQIARGLYLHBEAFVPIILHRLDKSNILLLEKIEHDDICNKTITDF 290
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 102 GTACD-IQTHMTNNKSGAAMWAEVFGSNYSKCDVFSWGIILWEVITRRKPFDEIGGP 160
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 291 GLAREWHTTKMSTAGTYANWAEVVKSSLFSGDSIWSGVLLWELLTGEVPIRGIDGL 350
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 161 AFPRIMWVHNGTRPPLIKNPKPIESLMTRCWSKDPQRSMEIEVKIMT 210
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 351 AVAYGVAVNKLTL-PIPTCTPEPPFAKLMEKCWQDDPHIRSPFALILEQLT 399
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
RESULT 6
US-09-862-027-18
; Sequence 18, Application US/09862027
; Patent No. US20020142428A1
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin R.
; TITLE OF INVENTION: No. US20020142428A1el Kinases and Uses Thereof
; FILE REFERENCE: 35800/234862
; CURRENT APPLICATION NUMBER: US/09/862,027
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 09/345,473
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 328
; TYPE: PRT
; ORGANISM: C. elegans
US-09-862-027-18

Query Match      27.6%; Score 346; DB 10; Length 328;
Best Local Similarity 34.9%; Pred. No. 8.6e-24;
Matches 76; Conservative 42; Mismatches 82; Indels 18; Gaps 8;

QY 2 ELRLSRVNHPIVVKLYGACLN--NPVCLVMEYAEAGSLYNVLHGAEPLPYTTAAHMS- 57
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 92 EAEILSKIRHKNITQFYGICKATGNDPFIIVTEVAEKGSLYDFTHSEBSQSFSASSGNSF 151
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 58 -----WCLQCSQGVAYLHNSQPKALIHRLDKPNNLLVAGTVLKICDFGTACDIQTHMT 112
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 152 DVVYKWSAQTSAGIQYLHYDAVDTHIHRDLKSNVNL-DKNLVCKICDFGTSKDL-THSC 209
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 113 NKK---GSAAMWAEV-FEGSNYSKCDVFSWGIILWEVITRRKPFDEIGGPAPRIMWAV 168
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 210 TAPSWGGAAMWSPEMLIQSEGLTATDVWSYGVLLWEILSKVEVPYKDY--EFRIPTWI 267
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 169 -HNGTRPPLIKNPKPIESLMTRCWSKDPQRSMEIE 205
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
; ORGANISM: Rattus norvegicus
US-09-947-199-8

Db 268 TQSGITLAIPPSPAPLKQLMSNCWKMTPKDRANMROI 305
; Sequence 2, Application US/09947199
; Patent No. US20020127684A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CARC PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MNI-068CP2
; CURRENT APPLICATION NUMBER: US/09/947,199
; CURRENT FILING DATE: 2001-09-05
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 60/111,938
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 09/458,457
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 835
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-947-199-2

Query Match      27.0%; Score 338.5; DB 10; Length 835;
Best Local Similarity 36.9%; Pred. No. 1.2e-22;
Matches 79; Conservative 42; Mismatches 78; Indels 15; Gaps 8;

QY 2 ELRLSRVNHPIVVKLYGACLN--PVCLVMEYAEAGSLYNVLHGAEPLPYTTAAHMSW 58
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 509 EVSILCQLNHPCVQFVGACLNDFSQFAIVTQYISGSLFSLH--FQKRILDQLSKLII 566
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 59 CLOCSQGVAYLHSM-QPKALIHRLDKPNNLLVAGTVLKICDFGTACDIQT----HMTN 113
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 567 AVDVAKGMEYLHNLTP--IIHRLDLSHNLLEYDGHAV-VADFGESRFLQSLDEDNMTK 623
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 114 NKGSAAMWAEV-FEGSNYSKCDVFSWGIILWEVITRRKPFDEIGGPAPRIMWAVHNGT 172
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 624 QFGNLRWMAPEVFTQCTRYTIKADVFSYALCLWEILTGEIPFAHLKPAADAAADWAYHH-I 682
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 173 RPPLIKNPKPIESLMTRCWSKDPQRSMEIEV 206
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 683 RPPIGYSIPKFISSLIRGNWACPEGRPEPSEVV 716
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
RESULT 8
US-09-947-199-8
; Sequence 8, Application US/09947199
; Patent No. US20020127684A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CARC PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MNI-068CP2
; CURRENT APPLICATION NUMBER: US/09/947,199
; CURRENT FILING DATE: 2001-09-05
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 60/111,938
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 09/291,839
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 835
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-947-199-8
```

[illegible]

```

      ; APPLICANT: Hannigan, Greg
      ; TITLE OF INVENTION: Intergrin-linked kinase and its Uses
      ; FILE REFERENCE: KIN-2CON
      ; CURRENT APPLICATION NUMBER: US/09/840,704
      ; CURRENT FILING DATE: 2001-04-23
      ; PRIOR APPLICATION NUMBER: 09/566,906
      ; PRIOR FILING DATE: 2000-05-09
      ; PRIOR APPLICATION NUMBER: US08/752,345
      ; PRIOR FILING DATE: 1996-11-19
      ; NUMBER OF SEQ ID NOS: 16
      ; SOFTWARE: FastSeq for Windows Version 4.0
      ; SEQ ID NO 5
      ; LENGTH: 263
      ; TYPE: PRT
      ; ORGANISM: H. sapiens
      ; FEATURE:
      ; NAME/KEY: Other
      ; LOCATION: (1)...(263)
      ;
US-09-840-704-5

Query Match          25.6%; Score 320; DB 10; Length 263;
Best Local Similarity 34.3%; Pred. No. 1,4e-21;
Matches 74; Conservative 43; Mismatches 91; Indels 8; Gaps 6;

OY    2 ELRQLSRNNHNIYKLYACALNP--VCLVMEYABEGSLYNVLHGAEPLPYTAAMASMC 59
      :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db    51 EVAIKKRLRHPIIVLFEMGAVTQPPNLSTVTEYLSRGSLYRLILHSGAREOLDERRRLSMA 110
      ::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
OY    60 LQCSGGVLYLMSMOPKALIHREDLKPNLLVAAGGVLCICDFG-TACDIOHTMTNKK--G 116
      ::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db    111 YDVAKGMNYTLNRNP-RIVHBDLKSPLLVDKITY-KVCPFGSLRAASTFLSSKSAAG 168
      ::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
OY    117 SAAMWAPFVGGSNYSEKCDYFSWGIIIMEVITRRKPDEIGGPAFRIMAWAHNGTRPPL 176
      :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db    169 TPENWAPFVLDEPSENKSDVYSFCVILLMELATIQGPNGNL-NPAQVAAVAGFKCKREI 227
      :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
OY    177 IKNLPKPIESLMTRCWSKDPSQRSPSMETVIKIMTHL 212
      :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db    228 PRNLNPQVAAIIEGCWTNEPKRPSPFATIMDLRLPL 263
      :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||:

RESULT 11
US-09-904-389-2
; Sequence 2, Application US/09904389
; Patent No. US20020129404A1
; GENERAL INFORMATION:
; APPLICANT: Clendenen, Stephanie K.
; APPLICANT: Schuster, Debra K.
; TITLE OF INVENTION: CTRI HOMOLOGUE FROM MELON
; FILE REFERENCE: 4257-0029.30
; CURRENT APPLICATION NUMBER: US/09/904,389
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: US 60/218,307
; PRIOR FILING DATE: 2000-07-14
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 850
; TYPE: PRT
; ORGANISM: Cucumis melo
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (154)...(154)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-904-389-2

Query Match          25.0%; Score 313.5; DB 10; Length 850;
Best Local Similarity 33.2%; Pred. No. 2e-20;
Matches 74; Conservative 47; Mismatches 89; Indels 13; Gaps 8;

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 10, 2002, 03:58:31 ; Search time 3101 Seconds

(without alignments)
2139.774 Million cell updates/sec

Title: US-09-830-144-2_COPY_76_303

Perfect score: 1252

Sequence: 1 VELRQLSRVNHPIVIXYGA.....MTHLMRYFFGADEPLQYPCQ 228

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% ,

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool/US09830144/runat_04122002_141754_4342/app_query.fasta_1.391
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=BLOSUM62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09830144 @CGN_1.1_3637/runat_04122002_141754_4342 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: .gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1252	100.0	1704	6	E38399 NF-kappa B
2	1252	100.0	1705	9	AB009358 Homo sapi
3	1252	100.0	1745	9	AF218074 Homo sapi
4	1252	100.0	2443	10	D76446 Mouse mRNA
5	1252	100.0	2757	9	BC017715 Homo sapi
6	1252	100.0	2769	6	AX377912 Sequence
7	1252	100.0	2769	9	AB009356 Homo sapi
8	1252	100.0	2785	6	E38397 NF-kappa B
9	1252	100.0	2850	9	AB009357 Homo sapi
10	1252	100.0	2866	6	E38398 NF-kappa B
11	1252	100.0	3107	10	BC006665 Mus muscu
12	1221	97.5	2812	5	U92030 Xenopus lae
13	1133	90.5	135147	2	AC114407 Mus muscu
14	700	55.9	3349	3	AF199466 Drosophil
15	700	55.9	3386	3	AF199466 Drosophil
16	599	47.8	32564	2	AC014558 Drosophil
17	599	47.8	164942	3	AC011758 Drosophil
18	599	47.8	302303	3	AE003571 Drosophil
19	392.5	31.3	3138	9	HSMSTM
20	392.5	31.3	3454	6	AX337846 Sequence
21	392.5	31.3	3454	9	HARNAMLK2
22	376	30.0	1445	5	AF265343 Danio rer
23	375.5	30.0	8125	3	AF481923 Dictyoste
24	374	29.9	1267	8	AY133876 Arabidops
25	374	29.9	1798	8	AY046026 Arabidops
26	373	29.8	1631	8	AY085535 Arabidops
27	373	29.8	1795	8	AY125513 Arabidops
28	372.5	29.8	3931	9	AF251442 Homo sapi
29	372	29.7	1429	10	AB049732 Mus muscu
30	372	29.7	3146	10	AB049731 Mus muscu
31	371	29.6	1365	6	AR119791 Sequence
32	371	29.6	1365	6	AR126751 Sequence
33	371	29.6	1365	6	AR128911 Sequence
34	371	29.6	1365	6	AR130842 Sequence
35	371	29.6	1365	6	AR138887 Sequence
36	371	29.6	1365	6	AR141355 Sequence
37	371	29.6	1368	9	AB049734 Homo sapi
38	371	29.6	1370	9	AF325454 Homo sapi
39	371	29.6	1571	9	BC001401 Homo sapi
40	371	29.6	2120	6	AR119790 Sequence
41	371	29.6	2120	6	AR126750 Sequence
42	371	29.6	2120	6	AR128910 Sequence
43	371	29.6	2120	6	AR130841 Sequence
44	371	29.6	2120	6	AR138886 Sequence
45	371	29.6	2120	6	AR141354 Sequence

ALIGNMENTS

RESULT 1

E38399 1704 bp DNA linear PAT 31-JAN-2002
LOCUS NF-kappa B activation inhibitory drug targeting TAK1 and method for
DEFINITION identifying the same.
E38399.1 GI:18626979
ACCESSION E38399.1
VERSION JP 2000197500-A/5.
KEYWORDS unclassified.
SOURCE unclassified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1704)
AUTHORS Sugita,N., Sakurai,H., Kageyama,N. and Hasegawa,H.
TITLE NF-kappa B activation inhibitory drug targeting TAK1 and method for
identifying the same
JOURNAL Patent: JP 2000197500-A 5 18-JUL-2000;
TANABE SEIYAKU CO LTD
COMMENT OS Unidentified
PN JP 2000197500-A/5
PD 18-JUL-2000
PF 04-FEB-1999 JP 1999026803
PR
PI NAOHISA SUGITA,HIROAKI SAKURAI,NORIKO KAGEYAMA, PI HIROSHI
HASEGAWA
PC C12Q1/48,A61K31/00,A61K31/00,A61K45/00,C12N5/10,C12N9/99, PC
C12Q1/02,
PC G01N33/15,G01N33/50,G01N33/566//C12N15/09,C12Q1/68,C12N15/09,
PC C12R1:91),
PC C12N5/00,C12N15/00,C12N15/00,C12R1:91)
CC Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1..1704
FT /organism='Unidentified'.
FEATURES
source location/Qualifiers
1..1704
/organism='unidentified'
/db_xref='taxon:32644'
BASE COUNT 511 a 381 c 401 g 411 t
ORIGIN
Alignment Scores: 3,266-117 Length: 1704
Pred. No.: 1252.00 Matches: 228
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB:
US-09-830-144-2_COPY_76_303 (1-228) x E38399 (1-1704)
QY 1 ValG1LeuA1rG1LeuSeArGValAaenH1sProAsn1LeVal1yLeuTyG1yAla 20
DB 226 GTAGACCTTCGGACGATTATCCCGTGGAACCATCCCAATATTGTAAGCTTATGAGCC 285
QY 21 CysLeuAsnProVal1yLeuVal1MetG1uTyTAAG1yG1ySeLeuTyrsenVal 40
DB 286 TGCTTAATCCAGTGTGTCTTGTGATGAGATATGCTGAAGGGGGCTTTTATATATG 345
QY 41 LeuH1G1yAlAG1uProLeuProTyTyTyTAAlA1H1sAlaMetSeT1rPcysLeu 60
DB 346 CTGCATGGTCTCAACCATATGCTATATATCTGCTGCCACGCCATGAGTTGGTCTTTA 405
QY 61 G1nCysSerG1nG1yAla1ATyLeuH1sSerMetG1nPro1yAlaLeu1LeH1sArG 80
DB 406 CAGTGTTCACCAAGAGTGGCTTATCTTCACAGATGCCAACCCAAAGCGCTAATTCACAG 465
QY 81 AspLeu1ySProProAsnLeuLeuLeuValAlAG1yG1yThyValLeu1yS1eCysAsp 100
DB 466 GACCTGAACACCAAACTTACTGCTGGTTCAGGGGGGCACTTCTTAAAAATTTGTGAT 525
QY 101 PhG1yThy1yAsp1yLeu1yG1yH1sMetThyAsnAsn1yG1ySer1yAla1ATP 120
DB 526 TTGTGATACACCTGTGACATTCAGACACATGACCAATTAACAGGGAGTGTCTGTGG 585

QY 121 MetAlaProG1uVal1PhG1yG1ySer1ySerG1u1yS1eCysAspVal1PhSeT1rP 140
DB 586 ATGACACCTGAAGTTTGTGAAGTATATACAGTGAAGAAATGACGCTCTTCAGCTGG 645
QY 141 G1y1Le1LeuTrG1uVal1LeThrArgArg1ySProPhAspG1u1LeG1yG1yPro 160
DB 646 GGTATATCTTGGGAATGATTAACGCTCGGAACCTTTGATGAGATTTGGTGGCCCA 705
QY 161 AlaPhArg1LeuTrpAlaVal1H1sAsnG1yThrArgProProLeu1Le1ySAsnLeu 180
DB 706 GCTTCCGATATATGAGGCTGTTCATATGATGATCGACCACTGATTAATAATTTA 765
QY 181 Pro1ySPro1LeuSer1eLeuMetThrArg1yS1rPSe1ySAspProSeG1nArgPro 200
DB 766 CCTTAAGCCCATGAGAGCTCGATGACTCGTTGTGTCTTAAGATCTTCCACAGGCCCT 825
QY 201 SerMetG1uG1u1LeVal1yS1eMetThrH1sLeuMetArgTyPheProG1yAlaAsp 220
DB 826 TCATGAGAGAAATTTGAAATATGACTCACTTGTATGCGTACTTTCAGAGACAGAT 885
QY 221 GluProLeuG1nTyPProCysG1n 228
DB 886 GACCACTTACAGTATCCTGTGAG 909
RESULT 2
AB009358 1705 bp mRNA linear PRI 20-JAN-2001
LOCUS Homo sapiens mRNA for TGF-beta activated kinase 1c, complete cdd.
DEFINITION AB009358
ACCESSION AB009358.2 GI:8978251
VERSION
KEYWORDS TAK1c; TGF-beta activated kinase 1c.
SOURCE Homo sapiens cell_line:Hela cDNA to mRNA, clone:pT7TAK1c.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (sites)
AUTHORS Sakurai,H., Shigemori,N., Hasegawa,K. and Sugita,T.
TITLE TGF-beta-activated kinase 1 stimulates NF-kappa B activation by an
NF-kappa B-inducing kinase-independent mechanism
JOURNAL Biochem. Biophys. Res. Commun. 243 (2), 545-549 (1998)
MEDLINE 98153801
JOURNAL 2 (sites)
REFERENCE Dempsey,C.E., Sakurai,H., Sugita,T. and Guesdon,F.
AUTHORS Alternative splicing and gene structure of the transforming growth
factor beta-activated kinase 1
JOURNAL Biochim. Biophys. Acta 1517 (1), 46-52 (2000)
MEDLINE 20568765
REFERENCE 3 (bases 1 to 1705)
AUTHORS Sakurai,H., Shigemori,N., Hasegawa,K., Sugita,T. and Guesdon,F.
TITLE Direct Submission
SUBMITTED (01-DEC-1997) Francois Guesdon, University of Sheffield,
Royal Hallamshire Hospital, Division of Molecular and Genetic
Medicine, Functional Genomics Group, Glossop road, Sheffield S10
2UF, United Kingdom (E-mail:f.a.guesdon@sheffield.ac.uk,
Fax:44-114-271-3846)
COMMENT On Jul 8, 2000 this sequence version replaced gi:2924627.
FEATURES
source location/Qualifiers
1..1705
/organism='Homo sapiens'
/db_xref='taxon:9606'
/clone='pT7TAK1c'
/translation='MSTASDAASSSSSSAGEMIEADSVLNFEEIDYKEIYEVEYGR
GAFGVCKAKRAKDVAIKQISESESRKAFFVELRQLSVNHNPVTKLIGACINPVCJ
VMEYVAGSGLYNVLHGAEPLPYTTAHAASWGCQSGQVAYVLSHMQPKALIHRLDKP
NLLVAVGTVLAKICDPTGACD1QTHMTNNKSAAMAAPEVFGSNYSKCDVFSWCI1

LOCUS AF218074.1 GI:6746614
DEFINITION Homo sapiens TGF beta-activated kinase splice variant d (TAK1)
mRNA, complete cds.
ACCESSION AF218074

BASE COUNT 511 a 381 c 402 g 411 t
ORIGIN

Alignment Scores:
Pred. No.: 3-26e-117 Length: 1705
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x AB009358 (1-1705)

QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIysLeuTyrGlyAla 20
Db 226 GTAGAGCTTCGGCAGTATCCCGTGTGAACCATCTTAATATTGTAAGCTTTATGGAGCC 285

QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyCysLeuTyrAsnVal 40
Db 286 TGCTTGAATCCAGTGTCTTGTGTGATGAATATGCTGAAGGGGGCTCTTTATATAATGTG 345

QY 41 LeuHisGlyAlaGluProLeuTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 60
Db 346 CTGCATGGTGCTGAACCATTTGCCATATTATACGTGCCACGCAATGAGTTGGTGTTTA 405

QY 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProIysAlaLeuIleHisArg 80
Db 406 CAGTGTTCCCAAGAGTGGCTTATCTTCACAGATGCACCCCAACGCTAATTCACAGG 465

QY 81 AspLeuIysProIleLeuLeuValAlaGlyGlyThrValLeuIysIleCysAsp 100
Db 466 GACCTGAAACCAACCACTTACTGCTGGTTGCAGGGGGACAGTCTTAAAAATTTGTGAT 525

QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnIysGlySerAlaAlaTrp 120
Db 526 TTTGGGTACAGCCCTGTGACATTCAGACACACATACCAATAACAGGGGAGTGTGCTGG 585

QY 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluIysCysAspValPheSerTrp 140
Db 586 ATGCGACCTGAAGCTTTTGAAGGTAGTAAATTAACAGTGAATAATGTACGCTTCAGCTGG 645

QY 141 GlyIleIleLeuTrpGluValIleThrArgArgGlyPheAspGluIleGlyGlyPro 160
Db 646 GGTATTATTCTTTGGGAAGTGATAACGCGTCGGAACCCCTTTTATGATGAGATTGGTGCCCA 705

QY 161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleIysAsnLeu 180
Db 706 GCTTTCCGAATCATGTGGCTGTTCATAATGGTACTCGACCCACCTGATAAAAAATTTA 765

QY 181 ProIysProIleGluSerLeuMetThrArgCysTrpSerIysAspProSerGlnArgPro 200
Db 766 CCTAAGCCCATTTGAGACCTGATGATCTCGTTGTGTCTAAAGATCCTTCCACGCGCCCT 825

QY 201 SerMetGluIleValIysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220
Db 826 TCAATGAGGAGAAATTTGAAAAAATAGTACTCTTGATGGGTACTTTCCAGGAGCAGAT 885

QY 221 GluProLeuGlnTyrProCysGln 228
Db 886 GAGCCATTACAGTATCCTTGTCTAG 909

RESULT 3
AF218074
LOCUS AF218074
DEFINITION Homo sapiens TGF beta-activated kinase splice variant d (TAK1)
mRNA, complete cds.
ACCESSION AF218074

VERSION AF218074.1 GI:6746614
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
gene
CDS
BASE COUNT 526 a 375 c 412 g 432 t
ORIGIN

Alignment Scores:
Pred. No.: 3-35e-117 Length: 1745
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x AF218074 (1-1745)

QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIysLeuTyrGlyAla 20
Db 232 GTAGAGCTTCGGCAGTATCCCGTGTGAACCATCTTAATATTGTAAGCTTTATGGAGCC 291

QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 40
Db 292 TGCTTGAATCCAGTGTGTCTTGTGATGAATATGCTGAAGGGGGCTCTTTATATAATGTG 351

QY 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 60
Db 352 CTGCATGGTGTCTGAACCATTTGCCATATTATCTGCTGCCACGCAATGAGTTGGTGTTTA 411

QY 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProIysAlaLeuIleHisArg 80
Db 412 CAGTGTTCCTCAAGAGTGGCTTATCTTCACAGATGCACCCCAACGCTAATTCACAGG 471

QY 81 AspLeuIysProProAsnLeuLeuValAlaGlyGlyThrValLeuIysIleCysAsp 100
Db 472 GACCTGAAACCAACCACTTACTGCTGGTTGCAGGGGGACAGTCTTAAAAATTTGTGAT 531

QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnIleGlySerAlaAlaTrp 120
DB 532 TTTGGTACAGCCCTGTGACATTCAGACACATGACCAATAAAGGGAGTGTCTGG 591
QY 121 MetAlProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 140
DB 592 ATGGCACCTGAACTTTTGAAGGTAGTAATACAGTGAATAATGTACCTTCCAGCTGG 651
QY 141 GlyIleIleuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 160
DB 652 GGTATTATTCTTTGGAGAGTAAACGGCTGGAAACCTTTGATGAGATTGTGGCCCA 711
QY 161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProleuIleLysAsnLeu 180
DB 712 GCTTTCCGATCATGTGGCTGTTCATTAATGTACTGCACCACTGATTAATAATTTTA 771
QY 181 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 200
DB 772 CTTAAAGCCATTAGAGCCTGATGACTCTGTGTGGTCTTAAGATCTTCCAGCGCCCT 831
QY 201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220
DB 832 TCAATGAGAGAAATGTGAAAATAATGACTCACTGATGCGGACTTTCAGAGAGAGAT 891
QY 221 GluProLeuGlnTyrProCysGln 228
DB 892 GAGCCATTACAGTATCTTGTCCAG 915
RESULT 4
MUSTAKI 2443 bp mRNA linear ROD 04-FEB-1999
LOCUS Mouse mRNA for TAK1 (TGF-beta-activated kinase), complete cds.
DEFINITION D76446.1 GI:1167505
VERSION TAK1; TGF-beta-activated kinase; protein kinase.
KEYWORDS Mus musculus cDNA to mRNA.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclirognathi; Muridae; Murinae; Mus.
REFERENCE 1 (sites)
AUTHORS Yamaguchi, K., Shirakabe, K., Shibuya, H., Irie, K., Oishi, I., Ueno, N.,
Taniguchi, T., Nishida, E. and Matsumoto, K.
TITLE Identification of a member of the MAPKKK family as a potential
mediator of TGF-beta signal transduction
JOURNAL Science 270 (5244), 2008-2011 (1995)
MEDLINE 96123277
REFERENCE 2 (bases 1 to 2443)
AUTHORS Matsumoto, K.
TITLE Direct Submision
JOURNAL Submitted (18-OCT-1995) Kunihiko Matsumoto, Faculty of Science,
Nagoya University, Department of Molecular Biology, Furoo-chou,
Chikusa-ku, Nagoya, Aichi 464-01, Japan
(E-mail: g44177a@nuc.cc.nagoya-u.ac.jp, Tel: 052-789-3000,
Fax: 052-789-3001)
FEATURES
source Location/Qualifiers
1..2443
/organism="Mus musculus"
/db_xref="taxon:10090"
157..1896
/codon_start=1
/product="TAK1 (TGF-beta-activated kinase)"
/protein_id="BA01184.1"
/db_xref="GI:1167506"
/translation="MSTASASSSSSSASSEMIEAPQVLNFEEDIKYEIEVEVGR
GAFGVCAKAKRAKDAIKOIESERAKAFIVEIROLRVRPNIVKLGACLPVCL
VMEVAGSLVNVHGAAPLPYTAHAMSMCLOCSQVAVLHMOPLALHRLKXP
NLLVAGSTVLKICDPGACDIQTHTMNNKSSAAMAEVSEGSNYSKCVESWGII
LMEVITRRKPPDEIGPARIMAVHNSTRPLIKNLKPIESLMTKWSKDPQORS
MEIVKIMTHLMRFPYGPADLPQYPCQYSDSGQSASATSGFMDIASNTNSKSDTN
MEOVPAITDITIKLESKLKNQAKQSSGSLSGASRGSVEBILPTSEKNSADM
SEIARIVATAGNQPRRSIODLIVTTEGQVSRSSPSVAMITTSGETSEKPAR
SHPTPDDSTDTNSDNSIPMAVILTLHDQLAPLPCPRSKSMVAFEDHCMAQEVYK

VOTETALLORKELEVALBEDQEDKQDQNTSRVQEHKKLLDENKSLSTYYQCKQKLE
VIRSQCKQCKGS"
BASE COUNT 669 a 568 c 647 g 559 t
ORIGIN
Alignment Scores:
Pred. No.: 4 87e-117 Length: 2443
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0
US-09-830-144-2_copy_76_303 (1-228) x MUSTAKI (1-2443)
QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20
DB 382 GTGAGGCTCCGGCAGTGTGTGCGGTGTAACCATCTTCAAGTTGTACAGGAGCC 441
QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 40
DB 442 TGCCGTGAATCCAGTATGCTTGTGATGGAATATGACAGAGGGGCTCATTTGATATGTC 501
QY 41 LeuHisGlyIleGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 60
DB 502 CTGCATGTGCTGAACCATTCCTTACTACACTGCTCTATGCACTGAGCTGTGTTA 561
QY 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80
DB 562 CAGTGTCCCAAGAGAGTGCGCTTACCTGCACAGCATGAGCCCAAGCGCTGATTACAGG 621
QY 81 AspleuLysProProAsnLeuLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100
DB 622 GACCTCAAGCCTCCAAACTTGCTGCTGGTGGCAGAGGAGCACTTCAAAAATCTCGAT 681
QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnIleGlySerAlaAlaTrp 120
DB 682 TTTGTACAGCTTGTGTGATCATCCAAACACATGACCAATATTAAGAGAGTGTGCTGG 741
QY 121 MetAlProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 140
DB 742 ATGGCGCTGAAGTGTGTAAGGAGCAATTACAGTGAAGTGTGCTTCACTGG 801
QY 141 GlyIleIleuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 160
DB 802 GGTATTATCTCTGGAGAGATGAACACGCGGAAACCTTCATGATGATCGGTGGCCCA 861
QY 161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProleuIleLysAsnLeu 180
DB 862 GCTTTCAGATCATGTGGCTGTTCATTAATGGCACTCGACCACTGATCAAAAATTTTA 921
QY 181 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 200
DB 922 CTTAAAGCCATTGAGACTGATGACACGCTGTGGTCTTAAGAACCATCTCAGCGCCCT 981
QY 201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220
DB 982 TCAATGAGAGAAATGTGAAAATAATGACTCACTGATGCGGACTTCCAGAGAGGAT 1041
QY 221 GluProLeuGlnTyrProCysGln 228
DB 1042 GAGCCATTACAGTATCTTGTCCAG 1065
RESULT 5
LOCUS BC017715 2757 bp mRNA linear PRI 06-DEC-2001
BC017715
DEFINITION Homo sapiens, mitogen-activated protein kinase kinase 7,
clone MGC:21263 IMAGE:3906857, mRNA, complete cds.
ACCESSION BC017715
VERSION BC017715.1 GI:17389342
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2757)
Strausberg, R.
Direct Submission
Submitted (03-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapps-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK plate: 22 Row: i Column: 18
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4507360.

FEATURES
source
1..2757
/organism="Homo sapiens"
/db_xref="LocustID:6885"
/db_xref="taxon:9606"
/clone="MGC:21263 IMAGE:3906837"
/issue_type="uterus, leiomyosarcoma"
/clone_lib="NIH MGC_71"
/lab_host="DH108"
/note="Vector: pCMV-SPORT6"
198..1937
/codon_start=1
/product="mitogen-activated protein kinase kinase 7"
/protein_id="AAH17715.1"
/db_xref="GI:17389343"
/translation="NSTGASASSSSSSAGEMIEAPSOVLNFEIDYKEIEVEVVGVR
GAGVVKAKWAKDAIKQIESERKAFIVELRLSRVNHNPVILVGLACLNPVCL
VMEYAGGSLYNVLHGAELPYTAAHMSWCLQSGVAYLHNSQPKALLIRDLKPP
NLLVAGGVLKIDFGTACDIQTHNNKSGAAWAEVEPGSNYSKCDVFSWGII
LMEVITRRPFDEIGGPAFRIMWAVHNGTRPELILNLPKPIESLMTKCSKDPQSRPS
MEEIVKIMTHLMRYFPGADEPLQPCQYDQGSNATSGSFMDIATNTSNKSDTN
MEQVPATNITIKRLSKLLKNOAKQSSGRLSLGASRGSSVESLPPTSEGRKMSADM
SEIAPRIATGNGQPRRSIODLTVTGPEQVSSRSRSPSRVMTTSGPTSEKPTR
SHPTPDDSDTDNGSDNSTPMAYLFDLHQLPLACPNSKSMVFEQHCWQAQBYMK
VTEPALLQRKQELVAELDDQKQNTSRVLVBHKKLLDNKSLSTYYQCKRQLE
VIRSQQQKQGTG"
BASE COUNT 819 a 566 c 647 g 725 t
ORIGIN

Alignment Scores:
Pred. No.: 5.57e-117 Length: 2757
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x BC017715 (1-2757)
Qy 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIysLeuTyrGlyVala 20
Db 423 GTAGAGCTTCGGCAGTATCCCGGTGAACCATCTTAATATTGTAAGCTTTATGAGCC 482
Qy 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnVal 40

Db 483 TGCTTGAATCCAGTGTGCTCTTGTGATGGAATATGCTGAAGGGGGCTCTTTATATAATGTG 542
Qy 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu 60
Db 543 CTGATGCTGCTGACCATTTGCCATATTATCTGCTGCCACGCAATGAGTGTGTTA 602
Qy 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80
Db 603 CAGTGTTCCTCCAAAGAGTGGCTTATCTTCACAGCATGCAACCCAAAGCGCTTAATTCACAGG 662
Qy 81 AspleuLysProProAsnLeuLeuValAlaGlyClyThrValLeuLysIleCysAsp 100
Db 663 GACCTGAAACCCACCAACTTACTGCTGTTGCGAGGGGGACAGTCTTAAATAATTTGTGAT 722
Qy 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyr 120
Db 723 TTTGGTACAGCTGTGACATTCACACACATGACCAATTAACAAGGGGAGTGTCTGTGG 782
Qy 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTyr 140
Db 783 ATGGCACCCTGAAGTCTTTTGAAGGTAGTAATTACAGTGAAAAATGTGACGTCTTCAGCTGG 842
Qy 141 GlyIleLeuLeuTyrGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 160
Db 843 GGTATTATCTTTGGGAAGTGATAACCGCTCGGAACCCCTTGTATGAGATTGGTGGCCCA 902
Qy 161 AlaPheArgIleMetTyrAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180
Db 903 GCTTTCGAATCATGTGGCTGTTCAATAATGTTACTGCACCACCATGATAAAAAATTTA 962
Qy 181 ProLysProIleGluSerLeuMetThrArgCysTyrSerLysAspProSerGlnArgPro 200
Db 963 CCTAAGCCCATTTGAGAGCCTGACTCGTTGTTGGTCTAAAGATCCTTCCACGCGCCT 1022
Qy 201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220
Db 1023 TCAATGGAGGAATTTGTAAAAATTAATGACTCACTTGTGCGGTACTTTCCAGGAGCAGAT 1082
Qy 221 GluProLeuGlnTyrProCysGln 228
Db 1083 GAGCCATTACAGTATCTCTTGTCAG 1106
RESULT 6
AX377912 AX377912 2769 bp DNA linear PAT 18-MAR-2002
LOCUS
DEFINITION Sequence 107 from Patent WO0212338.
ACCESSION AX377912
VERSION AX377912.1 GI:19573976
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS
TITLE Screening method
JOURNAL Patent: WO 0212338-A 107 14-FEB-2002;
Gruenthal GmbH (DE)
FEATURES
source Location/Qualifiers
1..2769
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 811 a 565 c 640 g 753 t
ORIGIN

Alignment Scores:
Pred. No.: 5.6e-117 Length: 2769
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

```

US-09-830-144-2_COPY_76_303 (1-228) x AX377912 (1-2769)
QY      1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIleuTyrgIyala 20
Db      388 GTAGAGCTTCGGCAGATTATCCCGTGGAGCAATCCCAATATTGTAAGCTTATGAGCC 447
QY      21 CysLeuAsnProValCysLeuValMetGluTYrAlaGluGlySerLeuTyraSnVal 40
Db      448 TGGCTTAATCTCAGTGTCTGTGTGATGAAATGCTGAAAGGGGGCTCTTATATATATG 507
QY      41 LeuHisGlyAlaGluProLeuProTyTYrTYrThraAlaAlaHisAlaMetSerTrpCysLeu 60
Db      508 CTGCATAGTGGTGAACCATTTGACATTTATCTGCTGCCACGCAATGATGGTGTGTTA 567
QY      61 GlnCysSerGlnGlyValAlaTYrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80
Db      568 CAGTGTTCACAGAGAGTGGCTTATCTTCACAGATGCCAACCCAAAGCGCTAATTCACAGG 627
QY      81 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThraValLeuLysIleCysAsp 100
Db      628 GACCTAAACACCAAACTTACTGCTGTGGCAGGGGGGACAGTCTTAAATAATTTGTGAT 687
QY      101 PheGlyThraAlaCysAspIleGlnThraHisMetThraAsnLysGlySerAlaAlaTrp 120
Db      688 TTTGGTACAGCCTGTGACATTCAGACACACATGACCAATACAGGGGAGTCTCTGTGG 747
QY      121 MetAlaProGluValPheGluGlySerAsnTYrSerGluLysCysAspValPheSerTrp 140
Db      748 ATGGCACCTGAAGTTTGAAGTGAATTAATACAGTGAATAATGACGCTTCACAGCTGG 807
QY      141 GlyIleIleLeuTrpGluValIleThraArgLysProPheAsnGluIleGlyGlyPro 160
Db      808 GGTATTATTCTTTGGAGAGTGAAGCGCTCGGAAACCTTTGATGAGATTGTGTGCCCA 867
QY      161 AlaPheArgIleMetTrpAlaValHisAsnGlyThraArgProProLeuIleLysAsnLeu 180
Db      868 GCTTCCGATCATGTGGCTGTTCATATGGTACTCCGACCACTGATTAATAAATTTA 927
QY      181 ProLysProIleGluSerLeuMetThraArgCysTrpSerLysAspProSerGlnArgPro 200
Db      928 CCTAAGCCCATGAGAGCCTGATGATGCTGTGTGTAAGATCTTCCACAGCGCCCT 987
QY      201 SerMetGluGluIleValLysIleMetThraHisLeuMetArgTYrPheProGlyAlaAsp 220
Db      988 TCAATGGAGAAATTTGTGAATAATGACTCACTTGATCGGTACTTTCAGAGCAAGAT 1047
QY      221 GluProLeuGlnTYrProCysGln 228
Db      1048 GAGCCATTACAGTATCTTGTCCAG 1071

RESULT 7
AB009356          2769 bp      mRNA      linear      PRI 04-MAR-1998
LOCUS            Homo sapiens mRNA for TGF-beta activated kinase 1a, complete cds.
DEFINITION      AB009356
ACCESSION      AB009356
VERSION      AB009356.1 GI:2924623
KEYWORDS      TAK1a; TGF-beta activated kinase 1a.
SOURCE      Homo sapiens lung cDNA to mRNA, clone_l1b:lambda gt11
              clone:PBSTAK1a.
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (sites)
AUTHORS      Sakurai,H., Shigemori,N., Hasegawa,K. and Sugita,T.
TITLE      TGF-beta-activated kinase 1 stimulates NF-kappa B activation by an
              NF-kappa B-inducing kinase-independent mechanism
JOURNAL      Biochem. Biophys. Res. Commun. 243 (2), 545-549 (1998)
MEDLINE      98153801
REFERENCE      2 (bases 1 to 2769)
AUTHORS      Sakurai,H.
TITLE      Direct Submision
JOURNAL      Submitted (01-DEC-1997) Hiroaki Sakurai, Lead Generation research
              laboratory, Tanabe Seiyaku Co. Ltd., 16-89, Kashima-3-chome,

```

```

Yodogawa-ku, Osaka, Osaka 532, Japan (E-mail:hsakurai@tanabe.co.jp,
Tel:++81 6 300 2571, Fax:++81 6 300 2593)
FEATURES
    source
        1..2769
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="PBSTAK1a"
            /tissue_type="lung"
            /clone_lib="Lambda gt11"
            163..1902
            /note="TAK1a"
            /product="TGF-beta activated kinase 1a"
            /protein_id="BA25025.1"
            /db_xref="GI:2924624"
            /translation="MSTASMASSSSSSAGEMIEAPQVLFNEEIDYEKEIEVEEYVGR
            GAFVYCKAKMPAKDVALKOISESEERKAFIYELKOLSVNHPNIVKLYGACLNIVCL
            VMEVAGGSLVNVLHGAEPLPYTTAAHMSMCLQCSQGVAYTHSMOPRALIHRLDKPP
            NLLVAGCTVCLKICDPGTACDIQTMTNKKSGAAMMAPEVEGSGSYSEKCDVFSNGITI
            LMEVITRRKPFDEIGPAPRIMVAHNGTRPPLIKLPEPIESLMTKWSKDPQSPS
            MEIVAKIMTHMKRYEPGADEPLQYPCQVSDQGSATSTGSFMDIATNTSNKSDTN
            MEQVPATNDITIKRLSKLKNQAKQOSESGRSLGASRGSVSESLPTSEGRMSADM
            SEIEARIAATGNGOPRRRSIQDLIVTGTPEQVSRSSSPSEVRMITTSGPTSEKPTR
            SHMPWDDSTDTNGSPNSIPMAVLTLDHQLAPCPNSKESMAVFEQHCMAQCYMK
            VQTEFALLQKQELVAELDQBEKQQTNSRLVQHKLIDENKSLSTTYQCKKQLE
            VTRSQQKQKQGS"
BASE COUNT      811 a      565 c      640 g      753 t
ORIGIN
Alignment Scores:
Pred. No.:      5,6e-117      Length:      2769
Score:          1282.00      Matches:      228
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:             9      Gaps:      0

US-09-830-144-2_COPY_76_303 (1-228) x AB009356 (1-2769)
QY      1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIleuTyrgIyala 20
Db      388 GTAGAGCTTCGGCAGATTATCCCGTGGAGCAATCCCAATATTGTAAGCTTATGAGCC 447
QY      21 CysLeuAsnProValCysLeuValMetGluTYrAlaGluGlySerLeuTyraSnVal 40
Db      448 TGGCTTAATCTCAGTGTCTGTGTGATGAAATGCTGAAAGGGGGCTCTTATATATATG 507
QY      41 LeuHisGlyAlaGluProLeuProTyTYrTYrThraAlaAlaHisAlaMetSerTrpCysLeu 60
Db      508 CTGCATAGTGGTGAACCATTTGACATTTATCTGCTGCCACGCAATGATGGTGTGTTA 567
QY      61 GlnCysSerGlnGlyValAlaTYrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80
Db      568 CAGTGTTCACAGAGAGTGGCTTATCTTCACAGATGCCAACCCAAAGCGCTAATTCACAGG 627
QY      81 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThraValLeuLysIleCysAsp 100
Db      628 GACCTGAACACCAAACTTACTGCTGTGGCAGGGGGGACAGTCTTAAATAATTTGTGAT 687
QY      101 PheGlyThraAlaCysAspIleGlnThraHisMetThraAsnLysGlySerAlaAlaTrp 120
Db      688 TTTGGTACAGCCTGTGACATTCAGACACACATGACCAATACAGGGGAGTCTCTGTGG 747
QY      121 MetAlaProGluValPheGluGlySerAsnTYrSerGluLysCysAspValPheSerTrp 140
Db      748 ATGGCACCTGAATTTTGAAGTGAATTAATCAGTGAATAATGATGATGAGCTTCACAGCTGG 807
QY      141 GlyIleIleLeuTrpGluValIleThraArgLysProPheAsnGluIleGlyGlyPro 160
Db      808 GGTATTATTCTTTGGAGAGTGAAGCGCTCGGAAACCTTTGATGAGATTGTGTGCCCA 867
QY      161 AlaPheArgIleMetTrpAlaValHisAsnGlyThraArgProProLeuIleLysAsnLeu 180

```

```

Db 868 GCTTCCGAATCATGTGGCGTGTTCATAATGGTACTCGACACACACTGATAAAAAATTTA 927
Qy 181 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 200
Db 928 CCTAAGCCATTGAGACCTGATGACTCGTGTGTGGTCTAAAGATCCTTCCAGCGCCCT 987
Qy 201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAasp 220
Db 988 TCAATGGAGGAAATGTGAAATAATGACTCCTTGATGGGTACTTTCCAGGAGCAGAT 1047
Qy 221 GluProLeuGlnTyrProCysGln 228
Db 1048 GAGCATTTACAGTATCCTTGTGCAG 1071

RESULT 8
E38397
LOCUS E38397 2785 bp DNA linear PAT 31-JAN-2002
DEFINITION NF-kappa B activation inhibitory drug targeting TAK1 and method for
identifying the same.
ACCESSION E38397.1 GI:18626977
VERSION JP 2000197500-A/3.
KEYWORDS JP 2000197500-A/3.
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 2785)
AUTHORS Sugita,N., Sakurai,H., Kageyama,N. and Hasegawa,H.
TITLE NF-kappa B activation inhibitory drug targeting TAK1 and method for
identifying the same
JOURNAL Patent: JP 2000197500-A 3 18-JUL-2000;
TANABE SEIYAKU CO LTD
OS Unidentified
PN JP 2000197500-A/3
PD 18-JUL-2000
PF 04-FEB-1999 JP 1999026803
PR
PI NAOHISA SUGITA,HIROAKI SAKURAI,NORIKO KAGEYAMA, PI HIROSHI
HASEGAWA
PC C12Q1/48,A61K31/00,A61K31/00,A61K45/00,C12N5/10,C12N9/99, PC
C12Q1/02.
PC G01N33/15,G01N33/50,G01N33/566//C12N15/09,C12Q1/68,(C12N15/09,
PC C12R1/91),
PC C12N5/00,C12N15/00,(C12N15/00,C12R1/91)
CC Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1..2785
FT Location/Qualifiers
source 1..2785
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 827 a 565 c 640 g 753 t
ORIGIN

Alignment Scores:
Pred. No.: 5.63e-117 Length: 2785
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x E38397 (1-2785)

Qy 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20
Db 388 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCATCCTAATATTGAAAGCTTTATGGAGCC 447
Qy 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 40
Db 448 TGCATTGAATCAGTGTCTTTGTGATGGAATATGCTGAAGGGGGCTCTTTATATAATGTG 507

```

```

Qy 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 60
Db 508 CTGCATGTGTGCTGAACCATTTGCCATATTATCTGCTGCCACGCAATGAGTTGGTGTTA 567
Qy 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80
Db 568 CAGTGTTCCTCCAGGAGTGGCTTATCTTCACAGCATGCAACCAAGCGCTAATTCACAGG 627
Qy 81 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100
Db 628 GACCTGAAACACCAACTTACTGCTGTTGCAGGGGGAGCAGTTCTAAAAATTTGTGAT 687
Qy 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 120
Db 688 TTTGGTACAGCCTGTGACATTCAGACACACATGACCAATAACAAGGGGAGTGTCTTGG 747
Qy 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 140
Db 748 ATGGCACCTGGAAGTTTTTTGAAGGTAGTAATTACAGTGAATAATGTGACGCTCTTCAGCTGG 807
Qy 141 GlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 160
Db 808 GGTATTATCTTTGGGAAGTAGTAACGCGTCGAAACCCCTTTGATGAGATTGGTGGCCCA 867
Qy 161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180
Db 868 GCTTTCCGAATCATGTGGGCTGTTCAATAATGTTACTCGACCACCATGATAAAAAATTTA 927
Qy 181 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 200
Db 928 CCTAAGCCCATTTGAGAGCCTGATGACTCGTGTGTGGTCTAAAGATCCTTCCAGCGCCCT 987
Qy 201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220
Db 988 TCAATGGAGGAAATGTGAAATAATGACTCCTTGATGGGTACTTTCCAGGAGCAGAT 1047
Qy 221 GluProLeuGlnTyrProCysGln 228
Db 1048 GAGCCATTACAGTATCCTTGTGCAG 1071

RESULT 9
AB009357
LOCUS AB009357 2850 bp mRNA linear PRI 04-MAR-1998
DEFINITION Homo sapiens mRNA for TGF-beta activated kinase 1b, complete cds.
ACCESSION AB009357
VERSION AB009357.1 GI:2924625
KEYWORDS TAK1b; TGF-beta activated kinase 1b.
SOURCE Homo sapiens lung cDNA to mRNA, clone_lib:lambda gt11
clone:pBSTAK1b.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Sakurai,H., Shigemori,N., Hasegawa,K. and Sugita,T.
TITLE TGF-beta-activated kinase 1 stimulates NF-kappa B activation by an
NF-kappa B-inducing kinase-independent mechanism
JOURNAL Biochem. Biophys. Res. Commun. 243 (2), 545-549 (1998)
MEDLINE 98153801
REFERENCE 2 (bases 1 to 2850)
AUTHORS Sakurai,H.
TITLE Direct Submission
JOURNAL Submitted (01-DEC-1997) Hiroaki Sakurai, Lead Generation research
Laboratory, Tanabe Seiyaku Co. Ltd.; 16-89, Kashima-3-chome,
Yodogawa-ku, Osaka, Osaka 532, Japan [E-mail:hsakurai@tanabe.co.jp,
Tel:++81 6 300 2571, Fax:++81 6 300 2593]
FEATURES
source 1..2850
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="pBSTAK1b"
/tissue_type="lung"
/clone_lib="Lambda gt11"
163..1983
CDS

```

```
/note="TAK1b"
/codon_start=1
/product="TGF-beta activated kinase 1b"
/protein_id="BAA25026.1"
/db_xref="GI:2924626"
/translation="MSTASASSSSSSAGEMIEAPSOVLNFEIDYKEIEVEVGR
GAPGVCKAKRAVDVAIKOIESESERAFVIEIROLSRVNHPITVLGYGCLNPVL
VMEVAREGSLYNVLHGAPPLPYTHAHAMSCLOCSQVATLHMOBKALIHDLKP
NLLVAGGTVLKICDFGTACDIQTHMTNKSAMMAVEVEGNSISEKCDVPSGII
LMEVITRRKPPDEIGGPAFRIMAVHNGTRPLIKLPEKIESIMTRCWSKDPQRS
MEELVKIMTHLMRYEPGADPELOYPCQYSDGQNSATISGFMASITMSKSDTN
MEQVATNDTTRKLESKIKNOAKOOSGSLISGASRGSESLPTSEKSKSAM
SEIRARIATAYASPKRKHRTASFGNIIILPVIISNGQRRRSIODIYVGTGP
GVSSRSSSPSVKMTTSGPTSEKTRSHPTPPDSTDITNSDMSIMAYITLDHQO
PLAPCNSKESNAVEFQHCMAQETMKVQITALLLOKQELVALLEODEKQDQNTSR
LVOEHKIKLDENKSLSTYYQCKQKQLEVIKRSQOKKROGTS"
```

```
BASE COUNT      832 a      588 c      656 g      774 t
ORIGIN
Alignment Scores:
Pred. No.:      5,786-117      Length:      2850
Score:      1252.00      Matches:      228
Percent Similarity:      100.00%      Conservative:      0
Best Local Similarity:      100.00%      Mismatches:      0
Query Match:      100.00%      Indels:      0
DB:      Gaps:      0
```

US-09-830-144-2_COPY_76_303 (1-228) x AB009357 (1-2850)

```
QY      1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIysLeuTyrglyAla 20
Db      388 GTAGAGCTTCGGAGTTATCCGTTGAAACCATCTTAATTGTAAGCTTATATGAGCC 447
QY      21 CysLeuAsnProValCysLeuValMetGluTyrglyGlySerLeuTyrglyAsnVal 40
Db      448 TGGTTGAATCCAGTGTCTGTTGATGAGATATGCTGAAGGGGCGCTTATATATATG 507
QY      41 LeuHisGlyAlaGluProLeuProTyryTyThrAlaAlaHisAlaMetSerTrpCysLeu 60
Db      508 CTGCATGCTGCTGAAACCATTTGCTTACCTGCTGCCACGCAATGATGGTGTTA 567
QY      61 GlnCysSerGlnGlyValAlaTyLeuHisSerMetGlnProIysAlaLeuIleHisArg 80
Db      628 GACCTGAACCAACCAACTTACTGCTGTTGACGGGGGACAGTTCTTAATAATTGTGAT 687
QY      101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnIysGlySerAlaAlaTrp 120
Db      688 TTTGGTACAGCCCTGTGACATTCGACACACATGACCAATTAACAGGGAGTCTCTTGG 747
QY      121 MetAlaProGluValPheGluGlySerAsnTyrySerGluIysCysAspValPheSerTrp 140
Db      748 ATGGCACTGGAAGTTTTTGAAGTACTTAATTACAGTGAATAATGTACCTTTCAGCTGG 807
QY      141 GlyIleIleLeuTrpGluValIleThrArgArgIysProPheAspGluIleGlyGlyPro 160
Db      808 CGATTAATCTTTGGAGAGTGAACGCGTCGGAACCCCTTGATGAGATTGTGTGGCCCA 867
QY      161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleIysAsnLeu 180
Db      868 GCTTTCCGAATCATGTGGGCTGTTCATAATGTGACTCGACCACTGATTAATAAATTTTA 927
QY      181 ProIysProIleGluSerLeuMetThrArgCysTrpSerIysAspProserGlnArgPro 200
Db      928 CTTAAACCCATTAGAGCCGTGATGACTCGTTGTGTTCTTAAGATCTTCCAGGCGCCCT 987
QY      201 SerMetGluGluIleValIysIleMetThrHisLeuMetArgTyryPheProGlyAlaAsp 220
Db      988 TCAATGAGAGAAATTGTGAATAATATGACTCACTTGAATGGTACTTTCAGAGAGAGAT 1047
```

```
QY      221 GluProLeuGlnTyryProCysGln 228
Db      1048 GAGCCATTACAGTATCTTGTCTCAG 1071
```

```
RESULT 10
E38398
LOCUS      E38398      2866 bp      DNA      linear      PAT 31-JAN-2002
DEFINITION      NP-kappa B activation inhibitory drug targeting TAK1 and method for
identifying the same.
```

```
ACCESSION      E38398
VERSION      E38398.1 GI:18626978
KEYWORDS      JP 2000197500-A/4.
SOURCE      unidentified.
ORGANISM      unidentified.
unclassified.
```

```
REFERENCE      1 (bases 1 to 2866)
AUTHORS      Sugita,N., Sakurai,H., Kageyama,N. and Hasegawa,H.
TITLES      NP-kappa B activation inhibitory drug targeting TAK1 and method for
identifying the same
Patent: JP 2000197500-A 4 18-JUL-2000;
```

```
JOURNAL      TANABE SEIYAKU CO LTD
COMMENT      OS      Unidentified
PN      JP 2000197500-A/4
PD      18-JUL-2000
PF      04-FEB-1999 JP 1999026803
PR
```

```
FI      NAOHISA SUGITA,HIROAKI SAKURAI,NORIKO KAGEYAMA, PI HIROSHI
HASEGAWA
PC      C12Q1/48,A61K31/00,A61K31/00,A61K45/00,C12N5/10,C12N9/99, PC
C12Q1/02,
PC      G01N33/15,G01N33/50,G01N33/566//C12N15/09,C12Q1/68,C12N15/09,
PC      C12R1.91),
PC      C12N5/00,C12N15/00,C12N15/00,C12R1.91)
CC      Strandedness: Double;
CC      Topology: Linear;
FH      Key
FT      source      1. 2866
FT      Location/Qualifiers
FT      source      1. 2866
FT      Location/Qualifiers
```

```
FEATURES
source      1. 2866
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT      848 a      588 c      656 g      774 t
ORIGIN
```

```
Alignment Scores:
Pred. No.:      5,826-117      Length:      2866
Score:      1252.00      Matches:      228
Percent Similarity:      100.00%      Conservative:      0
Best Local Similarity:      100.00%      Mismatches:      0
Query Match:      100.00%      Indels:      0
DB:      Gaps:      0
```

US-09-830-144-2_COPY_76_303 (1-228) x E38398 (1-2866)

```
QY      1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIysLeuTyrglyAla 20
Db      388 GTAGAGCTTCGGAGTTATCCGTTGAAACCATCTTAATTGTAAGCTTATATGAGCC 447
QY      21 CysLeuAsnProValCysLeuValMetGluTyrglyGlySerLeuTyrglyAsnVal 40
Db      448 TGGTTGAATCCAGTGTCTGTTGATGAGATATGCTGAAGGGGCTTATATATATG 507
QY      41 LeuHisGlyAlaGluProLeuProTyryTyThrAlaAlaHisAlaMetSerTrpCysLeu 60
Db      508 CTGCATGCTGCTGAAACCATTTGCTTACCTGCTGCCACGCAATGATGGTGTTA 567
QY      61 GlnCysSerGlnGlyValAlaTyLeuHisSerMetGlnProIysAlaLeuIleHisArg 80
Db      568 CAGTGTCCCAAGAGAGGCTTATCTTCACAGATGCAACCCCAAGCGCTTAATTCACAGG 627
QY      81 AspLeuIysProProAsnLeuLeuValAlaGlyIleThrValLeuIysIleCysAsp 100
```

```
Db 628 GACCTGAACACCAAACTTACTGCTGCTGGTTGCAGGGGGGACAGTCTTAAAAAATTTGTGAT 687
Qy 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrrp 120
Db 688 TTTGGTACAGCCTGTGACATTCAGACACACATGACCATAACCAAGGGAGTGTGCTTGG 747
Qy 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrrp 140
Db 748 ATGCACCTGAAGTTTGTGAAGGTAGTAATACAGTGAAAAATGTGACGTCTTCAGCTGG 807
Qy 141 GlyIleIleLeuTrrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 160
Db 808 GGTATTATCTTTGGGAAGTGATAACCGTCGGAACCCCTTGTATGAGATGTGTGGCCCA 867
Qy 161 AlaPheArgIleMetTrrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180
Db 868 GCITTTCCGAATCATGTGGCTGTTTCATAATGGTACTCGACACCACTGATATAAAATTTA 927
Qy 181 ProLysProIleGluSerLeuMetThrArgCysTrrpSerLysAspProSerGlnArgPro 200
Db 928 CCTAAGCCCATTTGAGAGCCTGATGACTCGTTGTTGGTCTAAAGATCCTTCCAGCGCCCT 987
Qy 201 SerMetGluGluIleValLysIleMetThrHisIleuMetArgTrrpPheProGlyAlaAsp 220
Db 988 TCAATGGAGGAAATGTGAAAAATAATGACTCACTTGATCGGTACTTTCCAGGAGCAGAT 1047
Qy 221 GluProLeuGlnTrrpProCysGln 228
Db 1048 GAGCATTTACGATATCTTTGTTCAG 1071
RESULT 11
BC006665 3107 bp mRNA linear ROD 07-AUG-2002
LOCUS Mus musculus, clone MGC:5989 IMAGE:3499247, mRNA, complete cds.
ACCESSION BC006665
VERSION BC006665.1 GI:13879375
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 3107)
Strausberg,R.
Direct Submission
Submitted (27-APR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunarathne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 9 Row: h Column: 9
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, GenomeScan gene prediction, Similarity but not identity
to protein.
Location/Qualifiers
1. .3107
/organism="Mus musculus"
```

```
/db_xref="taxon:10090"
/map="EVB/N"
/clone="MGC:5989 IMAGE:3499247"
/tissue_type="Mammary tumor, C3(1)-Tag model. Infiltrating
ductal Carcinoma, 5 month old virgin mouse."
/clone_lib="NCI_CGAP Mam6"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
148..1968
/codon_start=1
/product="Unknown (protein for MGC:5989)"
/protein_id="AAH06665.1"
/db_xref="GI:13879376"
/translaton="MSTASASSSSSSSASEMIEAPSOVLNFBEDIDYKEIEVEVVGR
GAFGVCKAKWRADKVAIKOIESEBKAFLVELROLSRVNHPIVKLYGACINPVCL
VMYAEAGSLYNVLHGAEPYPYTAHAMSKLCQSGVAYLHSMOPKALIHRLKPP
NLLEIVAGTGVKICDFTGACDIQTHMTNNKSAAMNAPEVFECSMTSEKCDVFSWGII
LWVITRRKPFDEIGGAPFAPRIMAVHNGTRPPLIKNLPKPIESLMTRCWSKDPSPRS
MEQVPATNDTIKRLESKLNKQAQSESLGASRGSVESLPSTSTNSKSDTN
SEIARIVATAAYSKPKRHKRTASFGNILDVPEIVISGNQPRRSIQDLTVTGTEP
GOVSSRSSPSRMITTSGPTSEKPARSHPEWDDSDTDNGSDNSIPMAYLIDHQLQ
FLAPCPNKSSEMAVFOHCKWAQEIATLALLOKQELVAELDDQDEKDOQNTSR
LVOEHKKLLDENKSLSTYYOQCKQKQEVIRSQOQRQGS"
BASE COUNT 898 a 687 c 772 g 750 t
ORIGIN
Alignment Scores:
Pred. No.: 6.36e-117 Length: 3107
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Dbs: 10 Gaps: 0
US-09-830-144-2_COPY_76_303 (1-228) x BC006665 (1-3107)
Qy 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTrrGlyAla 20
Db 373 GTGGAGCTCCGGCAGTGTCCGCTGTGAACCATCTTACATTGTCAGTTGACGAGCC 432
Qy 21 CysLeuAsnProValCysLeuValMetGluTrrAlaGluGlyGlySerLeuTrrAsnVal 40
Db 433 TGCCTGAATCCAGTATGCTTGTGATGGAATATGCAGAGGGGGCTCATTTGATAATGTG 492
Qy 41 LeuHisGlyAlaGluProLeuProTrrTrrThrAlaHisAlaMetSerTrrCysLeu 60
Db 493 CTGCATGTGTCTGAACCATTCCTTACTACACTGCTGCTCATGCCATGAGCTGTGTTTA 552
Qy 61 GlnCysSerGlnGlyValAlaTrrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80
Db 553 CAGTGTTCACAGGAGTGGCTTACTGACAGCATGACGCCAACGAGCGCTGATTCACAGG 612
Qy 81 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100
Db 613 GACCTCAGCCTCCAACTTCTGCTGCTGCAGAGGACAGTCTTAAAAAATCTGCCAT 672
Qy 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrrp 120
Db 673 TTTGTGTACAGCTTGTGACATCCAAACACACATGACCAATAATAAAGGAGTGTGCTTGG 732
Qy 121 MetAlaProGluValPheGluGlySerAsnTrrSerGluLysCysAspValPheSerTrrp 140
Db 733 ATGGCGCTGAAGTGTTTGAAGGTAGCAATTACAGTAAAAAGTGTGATGCTTCAGCTGG 792
Qy 141 GlyIleIleLeuTrrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 160
Db 793 GGTATTATCTCTGGGAAGTGTATTAACACCGCGGAACCCCTTCGATGAGATCGGTGGCCA 852
Qy 161 AlaPheArgIleMetTrrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180
Db 853 GCTTTTCAGAAATCATGTGGGCTGTTTCATAATGGCACTGCACCACTGATCAAAAATTTA 912
```

Qy	21	CysleuasnProValCysleuValMetGluThrAlaGluGlyGlySerLeuTyrAsnVal	40
Db	593	TGCCTAAATCCGGATGTTTATGATGGAGATATCCAGAGAGGGTTCAATGTATACAAAGTT	652
Qy	41	LeuHisGlyIaGluProLeuProTyrTyrThrAlaIaHisIaMetSerTyrCysleu	60
Db	653	TTGCATGAGAGCTGAACCGCTGCCCTTACATATACACTGCTCATGCAATGAGTGGTGTTTA	712
Qy	61	GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProGlySalLeuIleHisArg	80
Db	713	CAATGTGCCCAAGAGATGTGCATATTTACATGACATGAAAGCCGAAGGCTCTAATTCACAGG	772
Qy	81	AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValIleuLysIleCysAsp	100
Db	773	GACCTCAACCAACCAACCTGCTGCTGGTACTGGAGCACCGTTCTTAAACATTTGGTAC	832
Qy	101	PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyr	120
Db	833	TTTGGTACAGCCTGTGATATTCAACATCACATGACTAATACAAAGAAAGTCAGCATGG	892
Qy	121	MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTyr	140
Db	893	ATGGCTCCAGAAAGTTTTGGAAGGTAGCAATTACAGCGAAGAAATGTGACGTATTAGTTGG	952
Qy	141	GlyIleIleLeuTyrPgluValIleThrArgArgLysProPheAspGluIleGlyGlyPro	160
Db	953	GGAATAAATCTTTGGGAAGTATACCCGAGAAACCTTTCATGTAATGATGGTGCTCA	1012
Qy	161	AlaPheArgIleMetTyrPalaValHisAsnGlyThrArgProProLeuIleLysAsnLeu	180
Db	1013	GGCTTCCGATACATGTGGGCTGTTCACAATGTACTCGCCCAACATTATTTAAAAATTTG	1072
Qy	181	ProLysProIleGluSerLeuMetThrArgCysTyrSerLysAspProSerGlnArgPro	200
Db	1073	CCTAAGCCTATTTGAAGGTTATATGACTGCTGGTGGTCCAAAGATCCCCCAAGAACCT	1132
Qy	201	SerMetGluGluIleValLysIleMetThrHisLeuMetArgLysPheProGlyAlaAsp	220
Db	1133	TCAATGGAGGAGATGTGCACAAATATGACTCATCATATGACGATATTTTCCGGAGCGGAT	1192
Qy	221	GluProLeuGlnTyrProCysGln	228
Db	1193	GTTTCTTACAGATACCTTGTCAG	1216
RESULT 13			
AC114407			
LOCUS			
DEFINITION Mus musculus clone RP23-51G1, WORKING DRAFT SEQUENCE, 5 ordered			
pieces.			
AC114407			
VERSION AC114407.3 GI:21592111			
KEYWORDS HTG: HTGS PHASE2; HTGS_DRAFT; HTGS_FULLTOP.			
SOURCE house mouse.			
ORGANISM Mus musculus			
REFERENCE Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus.			
AUTHORS 1 (bases 1 to 135147)			
Birten,B., Linton,L., Nusbaum,C. and Lander,E.			
TITLE Mus musculus, clone RP23-51G1			
JOURNAL Unpublished			
REFA 2 (bases 1 to 135147)			
Birten,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,			
Anderson,S., Barua,N., Baselin,V., Bloom,T., Boguslavsky,L.,			
Boukhalaler,B., Brown,A., Camarata,J., Campobiano,A., Chang,J.,			
Chazaro,B., Chopel,Y., Colangelo,M., Collins,S., Collymore,A.,			
Cook,A., Cooke,P., Deatrelano,K., Dewar,K., Diaz,J.S., Dodge,S.,			
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galsgan,J., Gardyna,S.,			
Glade,S., Gordon,L., Goyette,M., Graham,L., Grand-Pierre,N.,			
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,			
Kamat,A., Karatas,A., Kells,C., Labrecque,K., Lamazares,R.,			
Landers,T., Lenoczky,D., Levine,R., Liu,G., Maclean,C.,			
Macedonia,P., Major,J., Margulis,N., Matthews,C., McCarthy,M.,			

McEwan, P., McKernan, K., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE

JOURNAL Submitted (08-MAR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE

AUTHORS

3 (bases 1 to 135147)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Illiev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrim, J., Meneus, L., Mihova, T., Milenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE

JOURNAL Submitted (26-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT

On Jun 26, 2002 this sequence version replaced gi:21536031.
 All repeats were identified using RepeatMasker:

Smith, A.F.P. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L23267

Center clone name: 51_G_1

----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 133686 bases at least Q40

Consensus quality: 134318 bases at least Q30

Consensus quality: 134777 bases at least Q20

Insert size: 151000; agarose-ftp

Insert size: 134747; sum-of-contigs

Quality coverage: 11.0 in Q20 bases; agarose-ftp

Quality coverage: 12.3 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently

* consists of 5 contigs. Gaps between the contigs

* are represented as runs of N. The order of the pieces

* is believed to be correct as given, however the sizes

* of the gaps between them are based on estimates that have

* provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 12654: contig of 12654 bp in length
 * 12655 12754: gap of 100 bp
 * 12755 13487: contig of 733 bp in length
 * 13488 13587: gap of 100 bp
 * 13588 14337: contig of 750 bp in length
 * 14338 14437: gap of 100 bp
 * 14438 63285: contig of 48848 bp in length
 * 63286 63385: gap of 100 bp
 * 63386 135147: contig of 71762 bp in length.
 FEATURES
 Location/Qualifiers
 1..135147
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="RP23-51G1"
 /clone_lib="RPCI-23 Female Mouse BAC"

source

1..12654
 /notes="assembly_fragment"
 clone_end:SP6
 vector_side:left"

misc_feature

12755..13487
 /notes="assembly_fragment"

misc_feature

13588..14337
 /notes="assembly_fragment"

misc_feature

14438..63285
 /notes="assembly_fragment"

misc_feature

63386..135147
 /notes="assembly_fragment"

misc_feature

clone_end:T7
 vector_side:right"

BASE COUNT

43081 a 25796 c 26087 g 39780 t 403 others

ORIGIN

Alignment Scores:
 Pred. No.: 4.79e-103 Length: 135147
 Score: 1133.00 Matches: 212
 Percent Similarity: 93.86% Conservative: 2
 Best Local Similarity: 92.98% Mismatches: 8
 Query Match: 90.50% Indels: 7
 DB: 2 Gaps: 2

US-09-830-144-2_COPY_76_303 (1-228) x AC114407 (1-135147)

Qy 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20

Db 22303 GTGAGAGCTCCGGCAGTTGTCACGTGTGAACCATCCTAATAATTGTAAGAGTTGTATGGAGCC 22362

Qy 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 40

Db 22363 TGCCTGCATCCAGTATGTTGTGATGGAATATACAGAGGGGGGCTCATTTGTAATGTG 22422

Qy 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu 60

Db 22423 CTGCATGTGTGTGAACCATTCGCTTACTACACTGCTCATGCAGTCAGCTGTGTGTTA 22482

Qy 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80

Db 22483 CAGCT-----TACCTGCACAGCATGCAGCCAAAGCTCTAATTCACAGG 22526

Qy 81 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100

Db 22527 GACCTCAAGCCTCCAAACTTCGCTGCTGTCAGAGGAGCAGTCTCAAAAATCTGGAT 22586

Qy 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyr 120

Db 22587 TTTGATACAGCTTGTGACATCCAGACACACATGACCAATAATAAAGGGAGTGTCTTGG 22646

Qy 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTyr 140

Db 22647 ATGGTCCCCGAAGTGTTTGAAGGTAGCAATTCACAGTAAAGAGTGTGTCTTCCAGCTGG 22706

Qy 141 GlyIleIleLeuTyrGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 160

Db 22707 GGTATTATCTCTGGGAAGTGATAACACGCTGGAACCC---GATGAGATCGTGGGCCCA 22763

QY 161 AAlapharGlIemetrPAlaValHIsaSnGlyThArGProProleuIleIlysAsnLeu 180
DB 22764 GCCTTCAGATTCAATGCGACTGTTCTTAATGAGCAGCCACCACTGATCAAAATTTA 22823
QY 181 ProIyProIleIleuSerIeumetThArGySTpSerIyAspProerGlnArGPro 200
DB 22824 CTTAAGCCCAATTCAGAGCTTGATGACACACTGTTGGTTAAGGATCTCTCAAGCCCT 22883
QY 201 SerMetGluGluIleValIySIIemetrThHIsleuMetArGyTyrPheProGlyAlaasp 220
DB 22884 TCATATGAGAAATTTGTGAATAATATGACTCACTTGATGCGGATCTCCACAGAGGGAT 22943
QY 221 GluProleuGlnItyrProCysGln 228
DB 22944 GAGCCATTACAGTATCTTGTCAG 22967
RESULT 14
AF199466 3349 bp mRNA linear INV 17-APR-2000
LOCUS Drosophila melanogaster TGF-beta activated-kinase 1 homolog mRNA,
DEFINITION complete cds.
ACCESSION AF199466
VERSION AF199466
KEYWORDS AF199466.1 GI:6289098
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 3349)
O'Connor,M.B., Shibuya,H. and Ueno,N.
TAKI participates in c-Jun N-terminal kinase signaling during
Drosophila development
Mol. Cell. Biol. 20 (9), 3015-3026 (2000)
20221548
JOURNAL
PUBMED 10757786
REFERENCE 2 (bases 1 to 3349)
Takahasu,Y., Nakamura,M., Stapleton,M., Danos,M., Matsumoto,K.,
O'Connor,M.B., Shibuya,H. and Ueno,N.
Direct Submision
Submitted (28-OCT-1999) Developmental Biology, National Institute
for Basic Biology, 38 Nishigonaka, Myodaiji-cho, Okazaki, Aichi
444-8585, Japan
FEATURES
source Location/Qualifiers
1..3349
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/chromosome="X"
/map="19E5-7"
916..2952
/note="TAKI homolog"
/codon_start=1
/product="TGF-beta activated-kinase 1 homolog"
/protein_id="AA06815.1"
/db_xref="GI:6289098"
/translation="MATASLDALQAAVVDSEITLRKVGHSYGVVCKAVWRDKLVA
VKEFASAEQDIEKEVQKSLRVPINIALHGYSVQOATYILMEABEGSLHNFH
GKVPAYSLAAMMARCAAGLAVLHAMTPKLIHRDVPILNLIKRNKICDF
GTVDKSTMTNNRNSGAAMWAPEVFGSKYTKEDIDISMAIIVMEVLSROPKGDIN
AYTIQMTYKGERPPLITTCPKRIEDLMTACWKTVPEDRSMQYIVAMHEIVKDTYG
ADKALETFVNOQIVKESDGTVAQPDSSSQGEUSPSTQLTPTTANAVNVAIA
ISKTTSTMENTSTSDITPTNSGQLDNNPLFYVWVNTKMDALPEBESNESNDSE
LTSAEATQRLLETIRNGWILMACPKMQLTLDVANGFDLSPSSSSSTNAKDSGE
RLTVDTQPVWMTTDLNNNGIGHANSNGLLSHANGQARDEILOEHEOELVNSID
VDVDEDENGTEOSLAEILIDPELOPEPPIPNABSOQLYRHRHAKRYLSVDNTL
YYADPRDKLIVQMDRTBRBQKOLIKMKDKBQQLYNNLOQVASYRDLAAGHHQ
PHPHPHQQLHPHSHPMHFLQDEGGCLPGSYGGSSESVBGMVVIPHNHA"
BASE COUNT 934 a 893 c 853 g 669 t
ORIGIN
Alignment Scores:

Pred. No.: 5,51e-61 Length: 3349
Score: 700.00 Matches: 130
Percent Similarity: 73.01% Conservative: 35
Best Local Similarity: 57.52% Mismatches: 5
Query Match: 55.91% Indels: 4
DB: 3 Gaps: 3
US-09-830-144-2_copy_76_303 (1-228) x AF199466 (1-3349)
QY 2 GluIeuArGlnIeuSerArGValAsnHIsProAsnIleValIySerIyGly----- 19
DB 1093 GAGGTGAAGCAAGTGTGCGCCGTAAGACCCGAACATCTGCTTCACGAGGATATCC 1152
QY 20 AlAcysLeuAsnProValCysIeuValMetGluTyrAlaGluIyGlySerIyTyrAsn 39
DB 1153 TCGTACAGACAGCCCACTCATATATGATGATGATGCGGAAAGTGATGCGTGCACAC 1212
QY 40 ValIeuHIsGlyAlaGluProIeuProIyTyrThrAlaAlaHIsAlaMetSerTyrCys 59
DB 1213 TTCCTTACAGGC--AAGGTGAAGCCGGCATATCTTCTGCGCCACCCCATGAGCTGGCG 1269
QY 60 LeuGlnCysSerGlnIyValAlaTyrLeuHIsSerMetGlnProIyAlaIeuIleHIs 79
DB 1270 CGCCATGTGCAAGGTCGTGCGCATTTTGCATGCCATAGCCCAACACCATTAATATCAT 1329
QY 80 ArgAspIeuIySProProAsnIeuLeuValAlaGlyGlyThrValIeuIySleCys 99
DB 1330 CGCAGCATGGAAGCGCGTGAACCTGCTTGAACCAAGAGGACGCAATCTGAAGATATGC 1389
QY 100 AspPheGlyThrAlaCysAspIleGlnThrHIsMetThrAsnAsnIySleIySerAlaAla 119
DB 1390 GACTTCGCGACGGTGGGGGCAAGTCGACCATGATATACCAACATGCGGTGTCGCT 1449
QY 120 TrpMetAlaProGluValPheGluIySerArGyTyrSerGluIyCysAspValPheSer 139
DB 1450 TGATGCGCGCCGACAGGCTTCGAAAGCTCAACTATACGAGAAAGTACATTTTCACG 1509
QY 140 TyrGlyIleIleIeuTyrGluValIleThrArgArgIySProPheAspIuIleGlyIy 159
DB 1510 TGGGCCATTTGTTCTATGAGGAGTTCGTGCAGAAACACCTTTAAAGCATGCACAT 1569
QY 160 ProAlaPheArGlnIemetrPAlaValHIsaSnGlyThArGProProleuIleIySAsn 179
DB 1570 --GCTAACCATTCAGATGAAGATCTACAGAGTGAAGCCGCCGCTGTGCACCACT 1626
QY 180 LeuProIySProIleIleuSerIeumetThArGySTpSerIyAspProserGlnArG 199
DB 1627 TGCCCCAAGCGCATTCAGACACTGATGACCGCTGTGAAAAAGTGGCCAGAGATGCC 1686
QY 200 ProSerMetGluIuIleValIySIIemetrThHIsleuMetArGyTyrPheProGlyAla 219
DB 1687 CCGTTCGATGACATATAGTGGGCGTTATGCAGAGATGTCGCAAGCATATACGGGGCG 1746
QY 220 AspGluProleuGlnItyr 225
DB 1747 GACAAAGCCCTGGAATAC 1764
RESULT 15
AY051953 3386 bp mRNA linear INV 27-AUG-2001
LOCUS Drosophila melanogaster LD42274 full length cDNA.
DEFINITION AY051953
ACCESSION AY051953
VERSION AY051953.1 GI:15292216
KEYWORDS FLI CDNA.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 3386)
Stapleton,M., Brokstein,P., Hong,L., Agrayani,A., Carlson,J.,
Champe,M., Chavez,C., Dorsett,V., Fatfan,D., Friese,E., George,R.,
Gonzalez,M., Guarin,H., Li,P., Liao,G., Miranda,A., Mungall,C.J.,

Nunoo,J., Pacieb,J., Paragas,V., Park,S., Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Rubin,G.M. and Celniker,S.
Direct Submission
Submitted (10-AUG-2001) Berkeley Drosophila Genome Project, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, USA
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory
Berkeley, CA 94720
This clone was sequenced as part of a high-throughput process to sequence clones from Drosophila Gene Collection 1 (Rubin et al., Science 2000). The sequence has been subjected to integrity checks for sequence accuracy, presence of a polyA tail and contiguity within 100 kb in the genome. Thus we believe the sequence to reflect accurately this particular cDNA clone. However, there are artifacts associated with the generation of cDNA clones that may have not been detected in our initial analyses such as internal priming, priming from contaminating genomic DNA, retained introns due to reverse transcription of unspliced precursor RNAs, and reverse transcriptase errors that result in single base changes. For further information about this sequence, including its location and relationship to other sequences, please visit our Web site (<http://fruitfly.berkeley.edu>) or send email to cdna@fruitfly.berkeley.edu.

FEATURES

source
1..3386
/organism="Drosophila melanogaster"
/strain="y; cn bw sp"
/db_xref="taxon:7227"
/map="19E1-19E1"
1..3386
/gene="CG18492"
/note="alignment with genomic scaffold AE003571"
/db_xref="FLYBASE:FBgn0042230"
935..2971
/gene="CG18492"
/notes="Longest ORF"
/codon_start=1
/product="LD4274p"
/protein_id="AAK93377.1"
/db_xref="GI:15292217"
/db_xref="FLYBASE:FBgn0042230"
/translation="WATASLDALQAYVDSEITREKVGHSYGVVCKAVMRDKLVA
VKEFPASARQDIKEVKOLSRVHPNIIALHGISYQOATVLI MEFAEGSLNHLFLH
GVKPAYSALDIAMSWARQAEGLAYLHAWPKPLIHRDVKPLNLLTNKGNLKICDF
GTVDKSTMTNNRGAAMAEVEFEGSKYTEKCDIFSWAIVLMEVLRSKQPFPGIDN
ADKALEYTVNQIIVTKESDGTVAAPDSLSQEGELSPSSQTLPPTTAANANVIA
ISKTTTSSMTNTSSDITPTNSGQDNNPLFYMTNWRDAIPEESNESRNDSPN
LTSSAPATQRLTIRNGMTLMACKPMBOILTDEANGFDLSPSSSSSTNAKSDGRE
RLVTDTKVMWTTDLNNGGIIHAHNSGLLSHANGQARDEELQEQHEQEIYNSLD
VVDPPEDENDTEQSLAILEDPEIQPEPIFNDASQLIYRDRHMAKRYLSVDVNL
YYAQDFKLIIVOMDTEREQKELLRMKQKGLQSLNNLQQOYASRLAAGHPQ
PHPHPNQLQHPHSHPPMHFTQDECGLLPGSVCGSGSEVBEGWVTPPHNA"

BASE COUNT 956 a 898 c 857 g 675 t
ORIGIN
Alignment Scores:
Pred. No.: 5.58e-61 Length: 3386
Score: 700.00 Matches: 130
Percent Similarity: 73.01% Conservative: 35
Best Local Similarity: 57.52% Mismatches: 57
Query Match: 55.91% Indels: 4
DB: 3 Gaps: 3

US-09-830-144-2_COPY_76_303 (1-228) x AY051953 (1-3386)

Qy 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGly----- 19
Db 1112 GAGGTGAAGCAGTGTCTCGCGGTGAAGACCCGACCATCATCTGCTCGACGGGATATCC 1171
Qy 20 AlaCysLeuAsnProValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsn 39

Db 1172 TCGTACCAGGCGCCACCTACCTGATAATGGAGTTCGCCGAAGGTGATCGCTGCACAAC 1231
Qy 40 ValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCys 59
Db 1232 TTCCTTCCACGCG---AAGGTGAAGCCGCGCATATCTCTGCCCCACGCCATGACGTGGCG 1288
Qy 60 LeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHis 79
Db 1289 CGCCAATGTGCAGAGGCTCTGGCATAATTTGCATGCGCATGACGCCAAAACCACTAATACAT 1348
Qy 80 ArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCys 99
Db 1349 CGCACGTGAAGCCGCTGAACCTCTTGACCAACAAGGCGCAATCTGGAAGATATGC 1408
Qy 100 AspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAla 119
Db 1409 GACTTCGGCAGCGTGGCGGCAAGTCAGCATGATGACCAACAATCGCGGAGTCCGCT 1468
Qy 120 TrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSer 139
Db 1469 TGGATGGCGCCGAGGCTTCGAAGGCTCCAAGTATACGGAGAAGTGTGACATTTTATGC 1528
Qy 140 TrpGlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGly 159
Db 1529 TGGGCCATTTGTTCTATGGAGGTTCTGTCCAGGAAGCAGCCCTTTAAAGGCATCGACAAT 1588
Qy 160 ProAlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsn 179
Db 1589 ---GCCTACACCATCCAGTGGAGATCTACAAGGTGAACGCCGCCGCTGCTGACCACT 1645
Qy 180 LeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArg 199
Db 1646 TGCCCCAAGCGCATCGAGGACCTGATGACCGCTGCTGGAACCGGTGCCGAGGATCGC 1705
Qy 200 ProSerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAla 219
Db 1706 CCGTCGATGCAGTACATAGTGGCGGTTATGACGAGATCGTCAAGGACTATACGGGGCG 1765
Qy 220 AspGluProLeuGlnTyr 225
Db 1766 GACAAGGCCCTTGAATAC 1783

Search completed: December 10, 2002, 04:57:50
Job time : 3132 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 10, 2002, 03:57:45 ; Search time 319 Seconds
(without alignments)
1609.579 Million cell updates/sec

Title: US-09-830-144-2_COPY_76_303

Perfect score: 1252

Sequence: 1 VELRLQLSRYNHPNIVKLYGA.....MTHLMRYFFGADEPLQYPCQ 228

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delopt 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool/US098301044/runat_04122002_141754_4332/app_query.fasta_1.391
-DB=N_geneseq_101002 -Qfmt=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US098301044 @CNG 1.1.79 @runat_04122002_141754_4332 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1252	100.0	1704	20	AAx99698	Human TGF-beta act
2	1252	100.0	1788	20	AAx56285	Human TAK1-6xHis e
3	1252	100.0	1959	18	AAT85095	Human transforming
4	1252	100.0	2443	18	AAT85094	Mouse transforming
5	1252	100.0	2656	20	AAx56279	Human TAK1 encodin
6	1252	100.0	2656	21	AAx39105	Human TAK-1 nucleo
7	1252	100.0	2769	24	ABL88437	Pain regulated cDN
8	1252	100.0	2785	20	AAx99696	Human TGF-beta act
9	1252	100.0	2866	20	AAx99697	Human TGF-beta act
10	700	55.9	3367	23	ABL02489	Drosophila melanog
11	599	47.8	10997	23	ABL02488	Drosophila melanog
12	414	33.1	759	23	ABL08337	Drosophila melanog
13	392.5	31.3	3454	24	ABL70018	Pancreas cancer re
14	374	29.9	1224	21	AAC43254	Arabidopsis thalia
15	373	29.8	1631	21	AAC39537	Arabidopsis thalia
16	371	29.6	1063	22	AAH34976	Human colon cancer
17	371	29.6	1365	21	AAZ99727	cDNA encoding huma
18	371	29.6	1706	21	AAx75674	DNA encoding a hum
19	371	29.6	2120	21	AAZ99726	cDNA encoding huma
20	371	29.6	2191	22	AAH99263	Human protein enco
21	371	29.6	2194	21	AAZ99734	Cardiovascular sys
22	371	29.6	2220	21	AAZ93783	Human survival reg
23	371	29.6	2254	21	AAZ99735	Cardiovascular sys
24	371	29.6	2403	22	AAx44701	Novel protein kina
25	371	29.6	2622	22	AAx75336	Human TGF-beta rec
26	371	29.6	3967	22	AAH73366	Human cervical can
27	370	29.6	2069	21	AAZ99737	Cardiovascular sys
28	370	29.6	2272	21	AAZ99736	Cardiovascular sys
29	363.5	29.0	3072	23	ABL04365	Drosophila melanog
30	361.5	28.9	3141	22	AAD18824	Human kinase (PKIN
31	361.5	28.9	3538	24	AAD34309	Human PKIN-12 CDNA
32	355.5	28.4	3066	24	ABQ86165	Novel human gene.
33	355	28.4	2157	22	AAH46913	cDNA encoding huma
34	355	28.4	3111	24	ABN86357	Novel human protei
35	355	28.4	3518	24	ABN86358	Novel human protei
36	353.5	28.2	3365	24	ABK84203	Human cDNA differe
37	353.5	28.2	3389	16	AAT01031	Human leucine-zipp
38	353.5	28.2	3389	18	AAH89349	Human leucine-zipp
39	352.5	28.2	3427	24	AB199250	Mouse ischaemic co
40	351.5	28.1	3109	23	ABL29755	Drosophila melanog
41	350	28.0	2283	21	AAC48526	Arabidopsis thalia
42	348	27.8	1591	21	AAC40839	Arabidopsis thalia
43	347	27.7	1659	21	AAC50138	Arabidopsis thalia
44	347	27.7	1661	21	AAC41921	Arabidopsis thalia
45	345.5	27.6	3558	24	ABK83874	Human cDNA differe

ALIGNMENTS

RESULT 1
ID AAX99698 standard; cDNA to mRNA; 1704 BP.
XX
AC AAX99698;
XX
DT 18-OCT-1999 (first entry)
XX
DE Human TGF-beta activated kinase (TAK) 1c encoding nucleotide sequence.
XX
KW Nuclear factor kappa B; NF-kB; inhibitor; TGF-beta activated kinase 1;
KW TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;
KW intractable disease; atrophic dermatitis; psoriasis; viral infection;
KW endotoxin shock; septicemia; human; hTAK1c; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers

FT CDS 1..1704
 FT /*tag= a
 FT /product= "hTAK1c"
 PN WO9940202-A1.
 PD 12-AUG-1999.
 PF 02-FEB-1999; 99WO-JP00422.
 PR 30-OCT-1998; 98JP-0309316.
 PR 06-FEB-1998; 98JP-0026003.
 XX (TANA) TANABE SEIYAKU CO.
 PA Hasegawa K, Kageyama N, Sakurai H, Sugita T;
 PI WPI: 1999-494298/41.
 DR P-PSDB; AAY28998.
 XX
 PT Nuclear factor kappa B activation inhibitors, useful as preventives
 for, e.g. autoimmune diseases
 XX
 PS Examples; Page 43-46; 49pp; Japanese.
 XX
 CC The invention provides a method for identifying or screening a nuclear
 CC factor kappa B (NF-kB) activation inhibitor by examining the effect of a
 CC test substance on modulation of the function(s) of TGF-beta activated kinase
 CC 1 (TAK1). The NFkB activation inhibitors targeting on TAK1 can be used to
 CC treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis),
 CC intractable diseases with inflammation (such as atrophic dermatitis and
 CC psoriasis), viral infection, endotoxin shock, septicemia and others. The
 CC present sequence represents the nucleotide sequence of human TAK1c
 CC (hTAK1c) protein.
 XX
 SQ Sequence 1704 BP; 511 A; 381 C; 401 G; 411 T; 0 other;

Alignment Scores:
 Pred. No.: 4.71e-133 Length: 1704
 Score: 1252.00 Matches: 228
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x AAX99698 (1-1704)

QY 1 ValGluLeuAArgGlnLeuSerArgValAsnHisProAsnIleValuLeuTyrGlyAla 20
 DB 226 GTAGAGCTTCGGCAGATTATCCCGTGAACCATCTTAATATTGTAAAGCTTATGGAGCC 285
 QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnVal 40
 DB 286 TGGTTGAATCCATGTCTCTTGTGATGGAATATGCTGAAGGGGGCTTTATATATATGG 345
 QY 41 LeuHisGlyValaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu 60
 DB 346 CTGCATGTGTCGTAACCATGTGATTAATCTGCTGCCACGACCATGAGTGGTGTGTTA 405
 QY 61 GlnCysSerGlnGlyValaAlaTyrLeuHisSerMetGlnProValaLeuIleHisArg 80
 DB 406 CAGTGTTCCTCCAGAGAGTGGCTTATCTTCACAGCATGCCAAACCGCTTAATTCACAGG 465
 QY 81 AspLeuLysProProAsnLeuLeuValaAlaGlyGlyThrValLeuLysIleCysAsp 100
 DB 466 GACCTGAACACCAACTACTGCTGGTGGCGAGGGGAGAGCTTTAAAAATTTGTGAT 525
 QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyr 120
 DB 526 TTGGTACACCTGTGACATTCAGACACACATGACCAATTAACAAAGGGAGTGGCTGCTGG 585
 QY 121 MetAlaProGluValaPheGluGlySerAsnTyrSerGlnLysCysAspAlaPheSerTyr 140

DB 586 ATGGCACCTGAAGTTTTTGAAGGAGTAATTACAGTGAATAAATGTGACGCTTCAGCTGG 645
 QY 141 GlyIleIleLeuTyrGluValaIleThrArgArgLysProPheAspGluIleGlyGlyPro 160
 DB 646 GGTATTATTCCTTTGGGAGGATGATACGCGTCGGAAAACTTTGATGATGATGGTGGCCCA 705
 QY 161 AlaPheArgIleMetTyrAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180
 DB 706 GCTTCCGAATCATGTGGGCTGTTCAATATGTAATGTAACACCATGATTAATAAATTTTA 765
 QY 181 ProLysProIleGluSerLeuMetThrArgCysTyrSerLysAspProSerGlnArgPro 200
 DB 766 CTTAAGCCCATTTGAGAGCCCTGATGACCTCGTTGGTCTAAAGATCCTCCAGGCCCTT 825
 QY 201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220
 DB 826 TCAATGAGGAAATTGTGAAAAATATGACTCACCTTGATGGGTATCTTCCAGAGCAGAT 885
 QY 221 GluProLeuGlnTyrProCysGln 228
 DB 886 GAGCCATTACAGTATCCTTGTCCAG 909

RESULT 2
 ID AAX56285 standard; DNA; 1788 BP.
 XX AAX56285
 AC AAX56285;
 XX
 DT 21-JUL-1999 (first entry)
 XX
 DE Human TAK1-6xHis encoding DNA.
 XX
 KW Human; TAB1; TAK1; screening; inhibition; TGF-beta;
 KM transforming growth factor beta; ss.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT CDS 7..1779
 FT /*tag= a
 PN WO9921010-A1.
 PD 29-APR-1999.
 PF 22-OCT-1998; 98WO-JP04796.
 PR 22-OCT-1997; 97JP-0290188.
 XX (CHUS) CHUGAI SEIYAKU KK.
 PA Ohnomo T, Ono K, Tsuchiya M;
 PI WPI: 1999-312645/26.
 DR P-PSDB; AAY09547.
 XX
 PT Screening for TGF- beta inhibitory substances, which are useful as
 PT drugs for treatment of diseases relating to its disorder
 XX
 PS Example 1; Page 167-171; 195pp; Japanese.
 XX
 CC A method has been developed for screening for substances which inhibit
 CC the binding of TAK1 polypeptide to TAB1 polypeptide. The method
 CC comprises: (a) contacting the polypeptide in the presence of a sample;
 CC and (b) detecting the amount of bound polypeptide, in which the sample
 CC can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming
 CC growth factor (TGF)-beta inhibitory substances can be used in drugs for
 CC indications e.g. as TGF-beta signal transduction inhibitors or
 CC activators, or extracellular matrix protein production enhancement
 CC inhibitors or activators, or cell proliferation prevention inhibitors or
 CC activators, or monocyte migration inhibitors or activators, or
 CC physiological activity induction inhibitors or activators, or

CC immunosuppression inhibitors or activators, or amyloid beta protein
CC precipitation inhibitors or activators, and such substances can also be
CC inhibitors of the TAK1 polypeptide function, particularly kinase
CC activity. The present sequence encodes TAK1-6xHis from an example of
CC the present invention.

XX Sequence 1788 BP; 551 A; 394 C; 413 G; 430 T; 0 other;
SQ

Alignment Scores:
Pred. No.: 5,04e-133 Length: 1788
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 20 Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x AAX56285 (1-1788)

QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20
Db 232 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCATCTCTAATATTGTAAAGCTTTATGGAGCC 291
QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 40
Db 292 TGTCTGAATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGGCTCTTTATATATGTG 351
QY 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 60
Db 352 CTGCATGTGTGCTGAACCATTTGTCATATTATCTGCTGCCCGCAATGATGTGGTGTTA 411
QY 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80
Db 412 CAGTGTCTCCAGGAGTGGCTTATCTTCACAGCATGCACCAACCAAGCGCTAATTCACAGG 471
QY 81 AspLeuLysProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100
Db 472 GACCTGAACACCACTTACTGCTGGTTCAGGGGGGACAGTCTTAAAAATTTGTGAT 531
QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 120
Db 532 TTTGGTACAGCTGTGACATTCAGACACATGACCATTAACAGGGGAGTCTGCTTGG 591
QY 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 140
Db 592 ATGGCACCTGAAGCTTTTGAAGGTAGTAAATTACAGTGAAAAATGTGACGTCTTCAGCTGG 651
QY 141 GlyIleIleLeuTrpGluValIleThrArgArgLysPhePheAspGluIleGlyPro 160
Db 652 GSPATATTCTTTGGGAAGTGATAACCGCTCGGAAACCCCTTTGATGAGATTGGTGGCCCA 711
QY 161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProLeuIleLysAsnLeu 180
Db 712 GCITTCGAAATCATGTGGGCTGTTCAATAGTGTACTCGACCACCTGATATAAAATTTA 771
QY 181 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 200
Db 772 CCTAAGCCCATTCAGAGCCTGATGACTCGTTGTTGTCTTAAAGATCCTTCCAGCGCCCT 831
QY 201 SerMetGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220
Db 832 TCAATGAGGAAATTTGAAAAATAATGACTCACTTGTATGGGTACTTTCCAGGAGCAGAT 891
QY 221 GluProLeuGlnTyrProCysGln 228
Db 892 GAGCCATTACAGTATCTCTTGTCTAG 915

RESULT 3

ID AAT85095
ID AAT85095 standard; cDNA; 1959 BP.

XX
AC AAT85095;

XX
DT 19-NOV-1997 (first entry)

XX Human transforming growth factor-beta activated kinase TAK-1 cDNA.
DE
XX TGF-beta; signal transduction; TGF-beta activated kinase;
KW MAPK kinase activator; AMK-1; bone morphogenetic protein; BMP;
KW protein kinase; ss.
XX
OS Homo sapiens.

XX Key Location/Qualifiers
FH CDS 183..1922
FT /*tag= a
FT /product= TAK-1
XX
XX JP09163990-A.
XX
XX 24-JUN-1997.

XX 27-SEP-1996; 96JP-0256747.
XX
XX 24-JUL-1996; 96US-0685625.
PR 29-SEP-1995; 95JP-0253549.
XX
XX (CHUS) CHUGAI PHARM CO LTD.
PA (UENO//) UENO N.

XX WPI; 1997-380171/35.
DR P-PSDB; AAW27093.
XX

PT DNA encoding transforming growth factor-beta-activated kinase, TAK-1
PT - useful for studying the TGF-beta signal transduction system
XX
XX Claim 9; Page 13-15; 20pp; Japanese.

XX The present sequence encodes human transforming growth factor-beta
CC (TGF-beta) activated kinase, TAK-1. The DNA is used to produce the
CC TAK-1 protein which is involved in the TGF-beta family signal
CC transduction system. TAK-1, also known as activator of MAPK Kinase
CC (AMK-1), is an enzyme which is activated by TGF-beta and bone
CC morphogenetic protein (BMP) and activates MAPK kinase by
CC phosphorylation.

XX SQ Sequence 1959 BP; 571 A; 447 C; 481 G; 460 T; 0 other;
SQ

Alignment Scores:
Pred. No.: 5,74e-133 Length: 1959
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x AAT85095 (1-1959)

QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20
Db 408 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCATCTCTAATATTGTAAAGCTTTATGGAGCC 467
QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 40
Db 468 TGCTTGAATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGGCTCTTTATATATGTG 527
QY 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 60
Db 528 CTGCATGTGTGCTGAACCATTTGCCCATATTATCTGCTGCCCGCAATGAGTGTGGTGTTA 587
QY 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80
Db 588 CAGTGTCTCCAGGAGTGGCTTATCTTCACAGCATGCACCAACCAAGCGCTAATTCACAGG 647
QY 81 AspLeuLysProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100
Db 648 GACCTGAACCACTTACTGCTGTTGACGGGGGACAGTCTTAAAAATTTGTGAT 707

QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaIleTrp 120
DB 708 TTTGGTACACCGCTGTGACATTCAGACACATGACCAATACAGGGAGGTGCTGCTGG 767
QY 121 MetAlaProGluValPheGluGlySerAsnTySerGluLysCysAspValPheSerTrp 140
DB 768 ATGGCACCTGAAGTTTGTGAAGTAGTAATTAACAGTGAATAATGTGACGCTTCAGCTGG 827
QY 141 GlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyPro 160
DB 828 GGTATTATTCTTTGGAGAGATACGCGTCGGAAACCTTTGATGAGATTGTGTGCCCA 887
QY 161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180
DB 888 GCTTTCGAAATCATGTGGGCTGTTCATATGTGACTCGACACCATGATTAATAATTTTA 947
QY 181 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 200
DB 948 CTTAAGCCCATTTGAGAGCCTGATGACTCGTTGTGGTTAAAGATCTTCCAGCGCCCT 1007
QY 201 SerMetGluGluIleValIleMetThrHisLeuMetArgTyRheProGlyAlaAsp 220
DB 1008 TCAATGAGGAAATTGTGAAATAATGACTCACTTGATCGGTACTTCCAGAGCAGAT 1067
QY 221 GluProLeuGlnTyRProCysGln 228
DB 1068 GAGCCATTACAGTATCTTGTCTG 1091
RESULT 4
AAT85094
ID AAT85094 standard; cDNA; 2443 BP.
XX
AC AAT85094;
XX
DT 19-NOV-1997 (first entry)
XX
DE Mouse transforming growth factor-beta activated kinase TAK-1 cDNA.
XX
KM TGF-beta; signal transduction; TGF-beta activated kinase;
KW MAPK kinase activator; AMK-1; bone morphogenetic protein; BMP;
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT CDS 157..1896
FT FT /*tag= a
FT FT /product= TAK-1
XX
PN JP09163990-A.
XX
PD 24-JUN-1997.
XX
PF 27-SEP-1996; 96JP-0256747.
XX
PR 24-JUL-1996; 96US-0685625.
XX
PR 29-SEP-1995; 95JP-0235349.
XX
PA (CHUS) CHUGAI PHARM CO LTD.
XX
PA (UENO) UENO N.
XX
DR MPI; 1997-380171/35.
DR P-PSDB; AAM27092.
XX
PT DNA encoding transforming growth factor-beta-activated kinase, TAK-1
XX - useful for studying the TGF-beta signal transduction system
XX
PS Claim 2; Page 10-12; 20pp; Japanese.
XX
CC The present sequence encodes mouse transforming growth factor-beta
CC (TGF-beta) activated kinase, TAK-1. The DNA is used to produce the
CC TAK-1 protein which is involved in the TGF-beta signal

CC transmission system. TAK-1, also known as activator of MAPK Kinase
CC (AMK-1), is an enzyme which is activated by TGF-beta and bone
CC morphogenetic protein (BMP) and activates MAPK kinase by
CC phosphorylation.
XX
SQ Sequence 2443 BP; 669 A; 567 C; 647 G; 559 T; 1 other;
Alignment Scores:
Pred. No.: 7 86e-133
Score: 1252.00 Length: 2443
Percent Similarity: 100.00% Matches: 228
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0
US-09-830-144-2_copy_76_303 (1-228) x AAT85094 (1-2443)
QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIleLeuTyRGIYAla 20
DB 382 GTGAGCTCCGGGAGTGTGCGGTGTGAACCATCTTAACATTGTCAAGTTGTACGAGCC 441
QY 21 CysLeuAsnProValCysLeuValMetGluTyRAlaGluGlySerLeuTyRAsnVal 40
DB 442 TGCCGTAATCCAGATGTCTTGATGATGATATGACAGAGGGGGCTCATTTGATATGTG 501
QY 41 LeuHisGluValGluProLeuProIleTyRThrAlaAlaHisIleMetSerTrpCysLeu 60
DB 502 CTGCATGTGTCTGAACCATTTGCTCTTACTACATGCTGCTCAATGAGCTGTGTTTA 561
QY 61 GlnCysSerGlnGlyValAlaTyRLeuHisSerMetGlnProLysAlaLeuIleHisArg 80
DB 562 CAGTGTTCACAGAGAGTGGCTTACCTGCACAGCATCAGCCCAACCGCTGATTCAGAG 621
QY 81 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100
DB 622 GACCTCAAGCCTCCAACTTGCTGTGTGTCAGAGGAGCAGTCTTAATAATCTGCAT 681
QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaIleTrp 120
DB 682 TTTGGTACACCTGTGTGACATTCACACACATGACCAATTAAGGAGGTGCTGCTGG 741
QY 121 MetAlaProGluValPheGluGlySerAsnTySerGluLysCysAspValPheSerTrp 140
DB 742 ATGGCGCCTGAAGTGTGAAAGTAGCAATTAACAGTGAAGTGTATGCTTCAGCTGG 801
QY 141 GlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyPro 160
DB 802 GGTATTATCTCTGGGAAAGTATACACGCCGGAACCTTCGATGAGATCGGTGCCCA 861
QY 161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180
DB 862 GCTTTCGAAATCATGTGGGCTGTTCATTAATGACATCGACACCATGATTAATAATTTTA 921
QY 181 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 200
DB 922 CTTAAGCCCATTTGAGAGCCTGATGACACGCTGTGTGTTAAGACCATCTCAGCGCCCT 981
QY 201 SerMetGluGluIleValIleMetThrHisLeuMetArgTyRheProGlyAlaAsp 220
DB 982 TCAATGAGGAAATTGTGAAATAATGACTCACTTATGCGGTATCCACAGAGGAGAT 1041
QY 221 GluProLeuGlnTyRProCysGln 228
DB 1042 GAGCCATTACAGTATCTTGTCTG 1065
RESULT 5
AAX56279
ID AAX56279 standard; DNA; 2656 BP.
XX
AC AAX56279;
XX
DT 21-JUL-1999 (first entry)
XX

DE Human TAK1 encoding DNA.

XX Human; TAB1; screening; inhibition; TGF-beta;
KW transforming growth factor beta; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
FH 183...1922
FT /*tag= a
FT

XX WO9921010-A1.

PN 29-APR-1999.

XX 22-OCT-1998; 98WO-JP04796.

XX 22-OCT-1997; 97JP-0290188.

XX (CHUS) CHUGAI SEIYAKU KK.

XX Ohtomo T, Ono K, Tsuchiya M;

XX WPI; 1999-312645/26.
DR P-PSDB; AAY09542.

PT Screening for TGF- beta inhibitory substances, which are useful as
PT drugs for treatment of diseases relating to its disorder

PS Example 1; Page 150-154; 195pp; Japanese.

XX A method has been developed for screening for substances which inhibit
CC the binding of TAK1 polypeptide to TAB1 polypeptide. The method
CC comprises: (a) contacting the polypeptide in the presence of a sample;
CC and (b) detecting the amount of bound polypeptide, in which the sample
CC can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming
CC growth factor (TGF)-beta inhibitory substances can be used in drugs for
CC indications e.g. as TGF-beta signal transduction inhibitors or
CC activators, or extracellular matrix protein production enhancement
CC inhibitors or activators, or cell proliferation prevention inhibitors or
CC activators, or monocyte migration inhibitors or activators, or
CC physiological activity induction inhibitors or activators, or
CC immunosuppression inhibitors or activators, or amyloid beta protein
CC precipitation inhibitors or activators, and such substances can also be
CC inhibitors of the TAK1 polypeptide function, particularly kinase
CC activity. The present sequence encodes human TAK1.

XX Sequence 2656 BP; 778 A; 557 C; 620 G; 701 T; 0 other;

SQ Alignment Scores:

Pred. No.: 8.85e-133 Length: 2656
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 20 Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x AAX56279 (1-2656)

QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20
Db 408 GTAGAGCTTCGGCAGTTATCCCGTGAACCATCTTAATATTGAAGCTTTATGGAGCC 467
QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyClySerLeuTyrAsnVal 40
Db 468 TGGTGAATCCAGTGTCTCTGTGTGAATATGCTGAAGGGGGCTCTTATATATATGTG 527
QY 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 60
Db 528 CTGCATGGTCTGAACCATTTGCCATATTATCTGCTGCCCAATGAGTTGGTGTATA 587
QY 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80

Db 588 CAGTGTTCCTCCCAAGAGTGGCTTATCTTTCACAGCATGCAACCCAAAGCGCTAATTCACAGG 647
QY 81 AspLeuLysProProAsnLeuLeuValAlaGlyClyThrValLeuLysIleCysAsp 100
Db 648 GACCTGAAACCAACCAACTTACTGCTGTTGCAGGGGGACAGTTCTAAAAATTTGTGAT 707
QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 120
Db 708 TTTGGTACAGCGCTGTGACATTCAGACACATGACCAATAACAAGGGAGTGTCTGCTTGG 767
QY 121 MetAlaProGluValPheGluGlySerAsnTyrSerCLeuLysCysAspValPheSerTrp 140
Db 768 ATGCACCTGAAGTTTGAAGGTAGTAATTACAGTGAAAAATGTGACGCTTTCAGCTGG 827
QY 141 GlyIleIleLeuTrpGluValIleThrArgLysProPheAspGluIleGlyGlyPro 160
Db 828 GGTATTATTCTTTGGGAAGTGATAACGCTCGGAACCCCTTTGATGAGATTGGTGGCCCA 887
QY 161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180
Db 888 GCTTTCCGAATCATGTGGCTGTTCAATAATGTTACTCGACCACCTGATAAAAAATTTA 947
QY 181 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 200
Db 948 CCTAAGCCCATTTGAGAGCCTGATGACTCGTTGTGGTCTAAAGATCCTTCCCAGCGCCT 1007
QY 201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220
Db 1008 TCAATGGAGGAATTGTGAAATAAATGACTCATTGATGGGTACTTTCAGGAGCAGAT 1067
QY 221 GluProLeuGlnTyrProCysGln 228
Db 1068 GAGCCATTACAGTATCTTGTGAG 1091
RESULT 6
AAA39105
ID AAA39105 standard; DNA; 2656 BP.
XX AAA39105;
AC AAA39105;
XX 04-SEP-2000 (first entry)
DT Human TAK-1 nucleotide sequence SEQ ID NO:1.
XX Human; TAK-1; TAB-1; mitogen activated protein kinase; MAPK;
DE screening; signal transduction; inhibition; inflammatory cytokine;
XX IL-1; interleukin 1; TNF; tumour necrosis factor; inflammation;
KW antiinflammatory; suppression; ds.
XX Homo sapiens.
OS
XX Key Location/Qualifiers
FH 183...1922
FT /*tag= a
FT /product= "TAK-1"
XX WO200023610-A1.
PN 27-APR-2000.
PD 21-OCT-1999; 99WO-JP05817.
PF 21-OCT-1998; 98JP-0299962.
PR (CHUS) CHUGAI SEIYAKU KK.
XX Tsuchiya M, Ohtomo T, Sugamata Y, Matsumoto K;
PI WPI; 2000-339707/29.
DR P-PSDB; AAY91000.
XX Method for screening inhibitors of TAK1 signal transduction for
PT suppression of inflammatory cytokine production and use as

PT antiinflammatory agents -
 XX
 PS Example 1; Page 73-80; 100pp; Japanese.
 XX
 CC The present invention describes a method for screening compounds for
 CC inhibition of inflammatory cytokine signal transduction by contacting
 CC the sample with TAK1 and its receptor TAB1 and selecting for inhibition
 CC of TAK1/TAB1 binding. Also described is a method for screening compounds
 CC for inhibition of inflammatory cytokine signal transduction in which the
 CC inhibition of TAK1 phosphorylation is selected for; and drug
 CC compositions for the treatment of inflammatory disorders containing as
 CC active component an inflammatory cytokine signal transduction inhibitor.
 CC TAK1 is an essential component of the signalling process which results
 CC in release of inflammatory cytokines such as interleukin-1 (IL-1),
 CC IL-10, tumour necrosis factor (TNF) and IL-6. The methods can be used
 CC for the selection of effective antiinflammatory agents. The present
 CC sequence encodes human TAK-1, which is used in the exemplification of the
 CC present invention.
 XX
 SQ Sequence 2656 BP; 778 A; 557 C; 620 G; 701 T; 0 other;
 Alignment Scores:
 Pred. No.: 8,856-133 Length: 2656
 Score: 1252.00 Matches: 228
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 21 Gaps: 0
 US-09-830-144-2_COPY_76_303 (1-228) x AAA39105 (1-2656)
 QY 1 ValGluLeuArgGlnLeuSerArgValaAsnHisProAsnIleValIlyLeuTyrgIyala 20
 DB 408 GTAGAGCTTCGGCAGATTATCCCGTGGAACCATCTTAATATTGTAAGCTTTATGAGACC 467
 QY 21 CysLeuAsnProValCysLeuValMetGluTyraIaGluGlySerLeuTyraAsnVal 40
 DB 468 TGGTTGAATCCAGTGTGCTTTGTGATGAAATATGCTGMAAGGGGGCTCTTTATATATGTG 527
 QY 41 LeuHisGlyAlaGluProLeuProTyTyTyThraAlaAlaHisAlaMetSerTyrcysLeu 60
 DB 528 CTGCATGCTGCTGAACCATTTGCCATATTATCTGCTGCCACGCAATGAGTGGTGTGTTA 587
 QY 61 GlnCysSerGlnGlyAlaAlaTyrlleuHisSerMetGlnProIyalaLeuIleHisArg 80
 DB 588 CAGGTGTTCCCAAGAGATGCTTATCTTCACACATGCAACCCMAAGCCGTAAATTTCACAG 647
 QY 81 AspLeuIyProProAsnLeuLeuValaIaGlyGlyThraValleuIySerCysAsp 100
 DB 648 GACCTGMAACCAACCAACTTACTGCTGTTGCAAGGGGGGACAGTTCTMAAATTGTGAT 707
 QY 101 PheGlyThraIaCysAspIleGlnThraHisMetThraAsnAsnGlySerAlaAlaTrp 120
 DB 708 TTTGGTACAGCGCTGTGACATTCAGACACACATGACCAATMAAGGGAGTGTCTGTGG 767
 QY 121 MetAlaProGluValaPheGluGlySerAsnTyTySerGluIyCysAspValPheSerTrp 140
 DB 768 ATGGCACCTGAAGTTTGAAGGTATTAACGTGAAGAAATGTGAGCTTCGACGTGG 827
 QY 141 GlyIleIleLeuTrpGluValaIleThraArgArgLysProPheAspGluIleGlyIyPro 160
 DB 828 GGTATTATTCTTTGGGAAGTGAATACGCGTCGMAAACCTTTGATGAGATTGGTGGCCCA 887
 QY 161 AlaPheArgIleMetTrpAlaValHisAenglyThraArgProProLeuIleLysAsnLeu 180
 DB 888 GCTTCCGATCATGTGGGCTGTTCATAATGTGTACTCGACCCACCTGATTAATAAATTTA 947
 QY 181 ProIyProIleGluSerLeuMetThraArgCysTrpSerIyAspProSerGlnArgPro 200
 DB 948 CCTAAGCCCATTAAGAGCCTGTATGACTCGTTGTGTCTTAAGATCCTTCCACGAGCCCT 1007
 QY 201 SerMetGluGluIleValIyIleMetThraHisLeuMetArgTyTyPheProGlyAlaAsp 220

DB 1008 TCATGAGAGCAATTGTCAAAATATATGACTTATGTCGGTACTTTCGAGAGAGAT 1067
 QY 221 GluProLeuGlnTyProCysGln 228
 DB 1068 GAGCCATTACGATATCTTGTGAG 1091
 RESULT 7
 ABL88437
 ID ABL88437 standard; cDNA; 2769 BP.
 XX
 AC ABL88437;
 DT 16-MAY-2002 (first entry)
 XX
 DE Pain regulated cDNA sequence 80.
 XX
 KW Pain; analgesic; gene therapy; neurological disorder;
 KW neurodegenerative disease; gene; ss.
 OS
 OS Homo sapiens.
 OS
 PN WO200212338-A2.
 XX
 PD 14-FEB-2002.
 XX
 PF 03-AUG-2001; 2001WO-EP09011.
 XX
 PR 03-AUG-2000; 2000DE-1037759.
 XX
 PA (CHEF) GRUENENTHAL GMBH.
 XX
 PI Gillen C, Wetzel S, Wenzel S, Wehne E, Schaefer MK;
 XX
 DR WPI; 2002-257469/30.
 DR P-PDB; ABB85033.
 XX
 PT Identifying pain-regulating compounds, useful for treating chronic pain
 PT and for diagnosis, by measuring binding of compounds to specific
 PT peptides and proteins -
 PS Claim 1; Fig 44; 213pp; German.
 XX
 CC The invention relates to identifying pain-regulating substances (A)
 CC comprises (1) incubating a test substance with a cell (or preparation
 CC from it) that has synthesised a peptide or protein (B) and (1i) measuring
 CC either binding of the test substance to (B) or some functional parameter
 CC that is altered by this binding. The method is useful for identifying
 CC pain-regulating substances (A) with analgesic activity. (A) along with
 CC nucleic acid (ABL88411-ABL88441) that encode proteins (B).
 CC ABB85006-ABB85037) that interact with (A); (B); vectors containing the
 CC nucleic acid; antibodies against (B); cells that express (B) and agents
 CC that bind to (B), are all useful for treating pain, particularly chronic
 CC pain, including use in gene therapy. The same materials can also be used
 CC for diagnosis e.g. of neurological and neurodegenerative diseases. The
 CC present sequence is that of a polynucleotide of the invention.
 XX
 SQ Sequence 2769 BP; 811 A; 565 C; 640 G; 753 T; 0 other;
 Alignment Scores:
 Pred. No.: 9,396-133 Length: 2769
 Score: 1252.00 Matches: 228
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0
 US-09-830-144-2_COPY_76_303 (1-228) x ABL88437 (1-2769)
 QY 1 ValGluLeuArgGlnLeuSerArgValaAsnHisProAsnIleValIlyLeuTyrgIyala 20
 DB 388 GTAGAGCTTCGGCAGATTATCCCGTGGAACCATCTTAATATTGTAAGCTTTATGAGACC 447
 QY 21 CysLeuAsnProValCysLeuValMetGluTyraIaGluGlySerLeuTyraAsnVal 40


```

Db 448 TGCTTGAATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGCTCTTTATATATATGTG 507
Qy 41 LeuHisGlyAlaGluProLeuProTyrThrAlaAlaHisAlaMetSerTrpCysLeu 60
Db 508 CTGCATGCTGCTGACCACTTGCATATATATCTGCTGCCCAATGAGTGTGTTTA 567
Qy 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuHisArg 80
Db 568 CAGTGTTCCTCCCAAGAGTGGCTTATCTTCACAGCATGCAACCCAAAGCGCTAATTCACAGG 627
Qy 81 AspleuLysProLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100
Db 628 GACCTGAAACCCAACTTACTGCTGTTGCGAGGGGGACAGTTCTAAAAATTTGTGAT 687
Qy 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 120
Db 688 TTTGGTACAGCCTGTGACATTCAGACACATGACCAATTAACAGGGGAGTGTGCTTGG 747
Qy 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 140
Db 748 ATGCGACCTGAAGTTTGAAGGTAGTAATTACAGTGAAAAATGTGACGTCTTCAGCTGG 807
Qy 141 GlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 160
Db 808 GGTATTATTCTTTGGGAAGTGATAACGCGTCGGAACCCCTTTGATGAGATTGTTGCGCCA 867
Qy 161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuLysAsnLeu 180
Db 868 GCITTTCCGAATCATGTGGGCTGTTCAATAGTACTCGACCACCATGATAAAAAATTTA 927
Qy 181 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 200
Db 928 CCTAAGCCCATGAGACCTGATGACTCGTTGTTGTCTTAAAGATCCTTCCAGCGCCCT 987
Qy 201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220
Db 988 TCAATGGAGGAATTTGAAAAATAATGACTCACTTGATGCGGTACTTTCCAGGAGCAGAT 1047
Qy 221 GluProLeuGlnTyrProCysGln 228
Db 1048 GAGCCATTACAGTATCTTGTTCAG 1071

RESULT 8
AAx99696
ID AAx99696 standard; cDNA to mRNA; 2785 BP.
XX
AC AAx99696;
XX
DT 18-OCT-1999 (first entry)
XX
DE Human TGF-beta activated kinase (TAK) 1a encoding nucleotide sequence.
XX
KW Nuclear factor kappa B; NF-kB; inhibitor; TGF-beta activated kinase 1;
KW TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;
KW intractable disease; atrophic dermatitis; psoriasis; viral infection;
KW endotoxin shock; septicemia; human; hTAK1a; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 163..1902
FT /tag= a
FT /product= "hTAK1a"
XX
XX WO9904202-A1.
XX
XX 12-AUG-1999.
XX
XX 02-FEB-1999; 99WO-JP00422.
XX
XX 30-OCT-1998; 98JP-0309316.
XX 06-FEB-1998; 98JP-0026003.

```

```

XX (TANA ) TANABE SEIYAKU CO.
XX Hasegawa K, Kageyama N, Sakurai H, Sugita T;
XX WPI; 1999-494298/41.
DR P-PSDB; AAY28996.
XX
PT Nuclear factor kappa B activation inhibitors, useful as preventives
PT for, e.g. autoimmune diseases
XX
PS Examples; Page 35-39; 49pp; Japanese.
XX
CC The invention provides a method for identifying or screening a nuclear
CC factor kappa B (NF-kB) activation inhibitor by examining the effect of a
CC test substance on modulating the function(s) of TGF-beta activated kinase
CC 1 (TAK1). The NFkB activation inhibitors targeting on TAK1 can be used to
CC treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis),
CC intractable diseases with inflammation (such as atrophic dermatitis and
CC psoriasis), viral infection, endotoxin shock, septicemia and others. The
CC present sequence represents the nucleotide sequence of human TAK1a
CC (hTAK1a) protein.
XX
SQ Sequence 2785 BP; 827 A; 565 C; 640 G; 753 T; 0 other;

```

```

Alignment Scores:
Pred. No.: 9.47e-133 Length: 2785
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 20 Gaps: 0

```

US-09-830-144-2_COPY_76_303 (1-228) x AAx99696 (1-2785)

```

Qy 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20
Db 388 GTAGAGCTCGCGCAGTTATCCCGTGTGAACCATCTCTAATATTGTAAAGCTTTATGGAGCC 447
Qy 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 40
Db 448 TGCITGAATCCAGTGTGCTTGTGATGGAATATGCTGAAGGGGGCTCTTTATATAATGTG 507
Qy 41 LeuHisGlyAlaGluProLeuProTyrThrAlaAlaHisAlaMetSerTrpCysLeu 60
Db 508 CTGCATGCTGCTGAACCATTCGCATATTTACTCTGCCACGCAATGAGTGTGTTTA 567
Qy 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuHisArg 80
Db 568 CAGTGTTCCTCCCAAGAGTGGCTTATCTTCACAGCATGCAACCCAAAGCGCTAATTCACAGG 627
Qy 81 AspleuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100
Db 628 GACCTGAAACCCAACTTACTGCTGTTGCGAGGGGGACAGTCTTAAAAATTTGTGAT 687
Qy 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 120
Db 688 TTTGGTACAGCCTGTGACATTCAGACACATGACCAATTAACAGGGGAGTGTGCTTGG 747
Qy 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 140
Db 748 ATGCGACCTGAAGTTTGAAGGTAGTAATTACAGTGAAAAATGTGACGTCTTCAGCTGG 807
Qy 141 GlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 160
Db 808 GGTATTATTCTTTGGGAAGTGATAACGCGTCGGAACCCCTTTGATGAGATTGTTGCGCCA 867
Qy 161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuLysAsnLeu 180
Db 868 GCITTTCCGAATCATGTGGGCTGTTCAATAGTACTCGACCACCATGATAAAAAATTTA 927
Qy 181 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 200

```

DB 928 CTTAAACCCATTGAGAGCCTGATGACTCGTTGTTGCTTAAGATCCTTCCAGGCGCCT 987
QY 201 SerMetGluGluIleValIysIleMetThrHisLeuMetArgTyrPheProGluAlaasp 220
DB 988 TCATATGAGGAAATGTGAAAATAATGACTCACTTGATCGGTACTTTCAGAGCAGAT 1047
QY 221 GluProLeuGlnTyrProCysGln 228
DB 1048 GAGCCATTACAGTATCCTTGTCTCAG 1071
RESULT 9
AA99697
ID AA99697 standard; cDNA to mRNA; 2866 BP.
XX
AC AA99697;
XX
DT 18-OCT-1999 (first entry)
XX
DE Human TGF-beta activated kinase (TAK) 1b encoding nucleotide sequence.
XX
KW Nuclear factor kappa B; NF-KB; inhibitor; TGF-beta activated kinase 1;
KW TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;
KW intractable disease; atrophic dermatitis; psoriasis; viral infection;
KW endotoxin shock; septicemia; human; hTAK1b; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 163..1983
FT /tag=a
FT /product="hTAK1b"
XX
PN WO9940202-A1.
XX
PD 12-AUG-1999.
XX
PF 02-FEB-1999; 99WO-JP00422.
XX
PR 30-OCT-1998; 98UP-0309316.
PR 06-FEB-1998; 98UP-0026003.
XX
XX
PA (TANA) TANABE SEIYAKU CO.
XX
PI Hasegawa K, Kagayama N, Sakurai H, Sugita T;
DR WPI; 1999-494298/41.
DR P-PSDB; AA928997.
XX
PT Nuclear factor kappa B activation inhibitors, useful as preventives
for, e.g. autoimmune diseases
XX
PS Examples; Page 39-43; 49pp; Japanese.
XX
CC The invention provides a method for identifying or screening a nuclear
CC factor kappa B (NF-KB) activation inhibitor by examining the effect of a
CC test substance on modulating the function(s) of TGF-beta activated kinase
CC 1 (TAK1). The NFkB activation inhibitors targeting on TAK1 can be used to
CC treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis),
CC intractable diseases with inflammation (such as atrophic dermatitis and
CC psoriasis), viral infection, endotoxin shock, septicemia and others. The
CC present sequence represents the nucleotide sequence of human TAK1b
CC (hTAK1b) protein.
XX
SQ Sequence 2866 BP; 848 A; 588 C; 656 G; 774 T; 0 other;

Alignment Scores:
Pred. No.: 9, 86e-133 Length: 2866
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 20 Gaps: 0

US-09-830-144-2_copy_76_303 (1-228) x AA99697 (1-2866)
QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIysIleuTyrGluAla 20
DB 388 GTACAGCTTGCGGCGATTATCCCGTGAACCATCTTAATATGTAAAGCTTTATGAGGCC 447
QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnVal 40
DB 448 TGCCTGAAATCCAGTGTCTTGATGAAATATGTGAAGGGGGCTTTATATTAATGTG 507
QY 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu 60
DB 508 CTGATGTGTCTGACCATTCATATTAATATATACGCTGCCACCATGATGATGCTTTTA 567
QY 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80
DB 568 CAGTGTTCACAGAGAGTGGCTTATCTTCAACGATGCACCAACCGCTAATTCACAG 627
QY 81 AspleuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100
DB 628 GACCTGAACACCAACCTTACTGTGTGGACGGGGGACAGTCTTAATAAATTTGTGAT 687
QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyr 120
DB 688 TTTGTATCAGCCTGTGACATTCAGACACATGACCAATTAACAGGGAGTCTGCTTGG 747
QY 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTyr 140
DB 748 ATGGCACCCTGAAGTTTGTGAAGGTAGTAATTAACAGTGAATAAAGTGAACGCTTCAC 807
QY 141 GlyIleIleLeuTyrGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 160
DB 808 GGTATTTATCTTGGGAAGATGATACCGCTGGAAACCTTTGATATAGATGGTGGCCCA 867
QY 161 AlaPheArgIleMetTyrAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180
DB 868 GCTTCCGAATCATGTGGCTGTTCATTAATGTAATCGACACACATCATTAATAAATTTTA 927
QY 181 ProLysProIleGluSerLeuMetThrArgCysTyrSerLysAspProSerGlnArgPro 200
DB 928 CTTAAGCCCATTTGAGAGCTGATGACTCGTTGTGGCTTAAGATCCTTCCAGGCCCT 987
QY 201 SerMetGluGluIleValIysIleMetThrHisLeuMetArgTyrPheProGluAlaasp 220
DB 988 TCATATGAGGAAATGTGAAAATAATGACTCACTTGATGCGGTACTTTCAGAGCAGAT 1047
QY 221 GluProLeuGlnTyrProCysGln 228
DB 1048 GAGCCATTACAGTATCCTTGTCTCAG 1071
RESULT 10
ABL02489
ID ABL02489 standard; cDNA; 3367 BP.
XX
AC ABL02489;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SHQ ID NO 1949.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
XX
XX 11-JUL-2000; 2000US-0614150.

```

XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX P-PSDB; ABB58386.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
XX Claim 1; SEQ ID NO 1949; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins
XX (ABB57737-ABB72072).
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 3367 BP; 938 A; 898 C; 856 G; 675 T; 0 other;

Alignment Scores:
Pred. No.: 1.77e-69 Length: 3367
Score: 700.00 Matches: 130
Percent Similarity: 73.01% Conservative: 35
Best Local Similarity: 57.52% Mismatches: 57
Query Match: 55.91% Indels: 4
DB: 23 Gaps: 3

US-09-830-144-2_COPY_76_303 (1-228) x ABL02489 (1-3367)

QY 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrgly----- 19
Db 1112 GAGTGAAGCAGTGTGCGCGGTGAAGCACCAGCAATCATCGCTTGCGCGGATATCC 1171
QY 20 AlaCysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsn 39
Db 1172 TCGTACCAGAGCCACCTACCTGATATGAGTTCGCCGAAGGTGGATGCTGCACAAAC 1231
QY 40 ValLeuHisGlyAlaGluProLeuProTyTyThrAlaAlaHisAlaMetSerTrpCys 59
Db 1232 TTCCTTCACGCG---AAGGTGAAGCGCGCATATCTCTGCGCCACGCCCATGAGCTGGCG 1288
QY 60 LeuGlnCysSerGlnGlyValAlaTyLeuHisSerMetGlnProLysAlaLeuIleHis 79
Db 1289 CGCCAATGTGCAGAGGTCTGGCATATTTCATGCCATGACGCAAAACCACTAATACAT 1348
QY 80 ArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCys 99
Db 1349 CGCGACGTGAAGCGCGTGAACCTGCTCTTGACCACAAAGGCGCAATCTGAAGATATGC 1408
QY 100 AspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAla 119
Db 1409 GACTTCGGCGACGFTGGCGGCAAGTCACCATGATGATGACCAACAATCGCGCAGTGCCT 1468
QY 120 TrpMetAlaProGluValPheGluGlySerAsnTyTrSerGluLysCysAspValPheSer 139
Db 1469 TGGATGGCGCCGAGGTCTTCGAAGGTCCAAGTATACGGAAGTGTGACATTTTAGC 1528
QY 140 TrpGlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGly 159
Db 1529 TGGGCCATTGTTCTTATGGAGGTTCTGCCAGGAAGCAGCCCTTTAAAGGCATCGACAAT 1588
QY 160 ProAlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsn 179
Db 1589 ---GCCTACACCATCCAGTGAAGATCTACAAGGGTGAACGCCCGCGCTGCTGACCACT 1645

```

```

QY 180 LeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArg 199
Db 1646 TGCCCCAAGCGCATCGAGGACCTGATGACCGCTGTGTGAAACGGTGTCCGAGGATCGC 1705
QY 200 ProSerMetGluGluIleValLysIleMetThrHisLeuMetArgTyThrPheProGlyAla 219
Db 1706 CCGTCGATGAGTACATAGTGGCGTTATGCACGAGATCGTCAAGGACTATACGGGGCG 1765
QY 220 AspGluProLeuGlnTyTr 225
Db 1766 GACAAGGCCCTTGAATAC 1783

RESULT 11
ABL02488
ID ABL02488 standard; cDNA; 10997 BP.
AC ABL02488;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 1946.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ss.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
XX
XX 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX P-PSDB; ABB58385.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
XX Claim 1; SEQ ID NO 1946; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins
XX (ABB57737-ABB72072).
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 10997 BP; 3237 A; 2362 C; 2382 G; 3016 T; 0 other;

Alignment Scores:
Pred. No.: 3.43e-57 Length: 10997
Score: 599.00 Matches: 130
Percent Similarity: 50.61% Conservative: 35
Best Local Similarity: 39.88% Mismatches: 57
Query Match: 47.84% Indels: 105
DB: 23 Gaps: 5

US-09-830-144-2_COPY_76_303 (1-228) x ABL02488 (1-10997)

```

```

QY      2  G1uLeuAArgGlnLeuSerArgValAsnHisProAsnHisLeuValLeuTyrGly----- 19
DB      2652  GAGGTGAACAGAGTTGTGTGGCGCGTGAAGCACCCGAACATCATCTCTGTGACCGGAGATTC 2711
QY      20  AlaCysLeuAsnProValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsn 39
DB      2712  TCGTACCAGCAGGCCACTACCTGATATAGAGTTGCGGAAGGTGATCGCTGCACAC 2771
QY      40  ValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCys 59
DB      2772  TTCCTTCACGGC---AAGGTGAAGCCGGCATATTTCTGTGCCACAGCCATGAGCTGGGCG 2828
QY      60  LeuGlnCysSerGln----- 64
DB      2829  CGCCATGTGCAGAGGTAGTGTCTGTGAGGGCGGGCGGTATTAACGTAGGGCTCA 2888
QY      64  ----- 64
DB      2889  CCCCCTAGCGCTTTCATCCAAATTCGTTAAATTATCTGTGAATAATGACTCAATG 2948
QY      64  ----- 64
DB      2949  ATTGTGTATTACATTCGTTCATGATTACTTAACAATTTCAATAGCTATCCGAT 3008
QY      64  ----- 64
DB      3009  AACCAAGATTAAATATATAGTAATTTGTTAAAGTTTAAAGCCCTTAATTCATTTCCATTG 3068
QY      65  -----GlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuHisAsr 80
DB      3069  TTTTCTCTCTAGGGGTCTGCGATTTTGTGATGTCATGACGCCAACAACACTATATCATCG 3128
QY      80  gAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysHisCysAs 100
DB      3129  CCACGTGAAGCCGCTGAACCTGCTCTTGACCAACAAGGACCGCATCTGAAGATGTGGA 3188
QY      100  pHeGlyThrAlaCysAspIleGlnThrHisMetThrAsnAnuLysGlySerAlaAlaTr 120
DB      3189  CTTTCGGACGGTGGCGGAACTGCATGATGACCAACAATCCGGCAGTGGCGCTTG 3248
QY      120  pMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTr 140
DB      3249  GATGGCGGCCGAGGTCTTGAAGGCTCCACAGTATACGGAAGTGTGACATTTTACGTG 3308
QY      140  pGlyIleIleLeuTrrpGluValIleThrArgArgLysProPheAspGluIleGlyLyr 160
DB      3309  GGCCATTGTTCTATGGAGGTCTGTCCAGGAAGCAGCCCTTTAAAGCATCGACAAAT-- 3366
QY      160  oAlaPheArgIleMetTrpAlaValHisAspGlyThr----- 172
DB      3367  -GCCTAACCAATCCAGTGAAGATCTTACAAAGGTTGC-GTCTCCCAATTCACCTTTTTC 3424
QY      173  -----ArgProLeuIleLysAsnLe 180
DB      3425  ATCGAGCTTATCGAGATCTGTGCTCTCGCAGGTGAAGCCCGCGCTGACACACTTG 3484
QY      180  uProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnAsp 200
DB      3485  CCCCAGGCGCATCGAGACCTGATGACCGCGCTGGGAAAAAGGTGGCCGAGATCGGCC 3544
QY      200  oSerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAs 220
DB      3545  GTTCATGCAATGATAGTGGGGGTATGACAGAGATCTGTCAAGGACTATATACGGGGCGGA 3604
QY      220  pGluProLeuGlnTyr 225
DB      3605  CAAGGCCCTGGAATAC 3620
RESULT 12
ABL08337
ID      ABL08337 standard; cDNA; 759 BP.
XX

```

```

AC      ABL08337;
XX
XX      26-MAR-2002 (first entry)
DT
XX
DE      Drosophila melanogaster expressed polynucleotide SEQ ID NO 19493.
KW      Drosophila; developmental biology; cell signalling; insecticide;
KW      pharmaceutical; gene; ss.
XX
OS      Drosophila melanogaster.
XX
PN      MO200171042-A2.
XX
PD      27-SEP-2001.
XX
PF      23-MAR-2001; 2001WO-US09231.
XX
PR      23-MAR-2000; 2000US-191637P.
PR      11-JUL-2000; 2000US-0614150.
XX
PA      (PEKE ) PE CORP NY.
PI      Venter JC, Adams M, Li PWD, Myers EW;
PI      MPI; 2001-656860/75.
DR      P-PDB; ABB64234.
XX
XX      New isolated nucleic acid detection reagent for detecting 1000 or more
PT      genes from Drosophila and for elucidating cell signalling and cell-cell
PT      interactions -
PS      Claim 1; SEQ ID NO 19493; 21bp + Sequence listing; English.
XX
XX      The invention relates to an isolated nucleic acid detection reagent
CC      capable of detecting 1000 or more genes from Drosophila. The invention is
CC      useful in developmental biology and in elucidating cell signalling and
CC      cell-cell interactions in higher eukaryotes for the development of
CC      insecticides, therapeutics and pharmaceutical drugs. The invention
CC      discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC      sequences (ABL161840-ABL16175) and the encoded proteins
CC      (ABBS7737-ABBS7072).
CC      The sequence data for this patent did not form part of the printed
CC      specification, but was obtained in electronic format directly from WIPO
CC      at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ      Sequence 759 BP; 226 A; 162 C; 198 G; 173 T; 0 other;
XX
Alignment Scores:
Pred. No.: 1,126-37 Length: 759
Score: 414.00 Matches: 82
Percent Similarity: 59.69% Conservative: 35
Best Local Similarity: 41.84% Mismatches: 53
Query Match: 33.07% Indels: 26
DB: 23 Gaps: 4
XX
US-09-830-144-2_COPY_76_303 (1-228) x ABL08337 (1-759)
QY      2  G1uLeuAArgGlnLeuSerArgValAsnHisProAsnHisLeuValLeuTyrGlyAlaCys 21
DB      160  GAGATCTACAGGTGACCAAGGCCAGCATGTCAACATAGTTGACTTACGGCACATCG 219
QY      22  LeuAsnProValCys-----LeuValMetGluTyrAlaGluGlySerLeuTyrAsn 39
DB      220  AGGACAGAGGATGCGCCCTGTGTGATGGAATTCGTAAGCGGTGATCTGTGCCAGT 279
QY      40  ValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCys 59
DB      280  TTTCTGCAC---GCGAAAGCAAGCCAAAGTTATGTGCATGCCACGCTTCAACTGGGCG 336
QY      60  LeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuHis 79
DB      337  CATGAGATGCTCAGGGCATACCTATCTGATGCGCATGCGACCGAAGCAAGTATTCAT 396

```

```
Qy 80 ArgAspLeuLysProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCys 99
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 397 CCGGATATAAGCCCACTCAATACACTGCTATCGGAGAGGACTCAACTGAGATTGC 456
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 100 AspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAla 119
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 457 GATTTCCGAACCTGTTGTGGACCTATCCCAATCGATATCGTGAATGCGGGCACCTGCAGA 516
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 120 TrpMetAlaProGlu----- 124
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 517 TACAAGCGCCCGAGGTAAGGGAGTGTGTTGATTTCAATCAATCGAATAATATCAAC 576
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 125 -----ValPheGluGlySerAsnTyrSerGluLysCysAspVal 137
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 577 CAACCAACCGGCTTTCAAAAGTTCTACAAGGAATAAACCCGATGAAAGTGGCATGTG 636
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 138 PheSerTrpGlyIleIleLeuTrpGluValIleThrArgLysProPheAspGluIle 157
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 637 TATAGTTGGCTATTACCTTTTGGGAATATTGTGCGCAAGGAGCCATTTCAGCAATAT 696
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 158 GlyGlyProAlaPheArgIleMetTrpAlaValHisAsnGlyThrArg 173
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 697 ---AATACGCTTTTGAACCTATACATGCGCTATTATGAAGGCAAGAGA 741
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
RESULT 13
ABL70018
ID ABL70018 standard; DNA; 3454 BP.
XX
AC ABL70018;
XX
DT 15-MAY-2002 (first entry)
XX
DE Pancreas cancer related gene sequence SEQ ID NO:8355.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-234927P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
```

```
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 02-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX
XX (AVAL-) AVALON PHARM.
XX
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
XX Soppet DR, Weaver Z;
XX WPI; 2002-188264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a
XX chemical agent to be tested for anti-neoplastic activity, and
XX determining a change in expression of a gene of a signature gene set -
XX
XX Claim 1; SEQ ID 8355; 44pp; English.
XX
XX The present invention describes a method (M1) for screening for an
XX anti-neoplastic agent. The method involves exposing cells to a chemical
XX agent to be tested for anti-neoplastic activity, determining a change in
XX expression of at least one gene (I) of a signature gene set, where (I)
XX comprises a sequence (S) selected from 8447 sequences (given in ABL61664
XX to ABL70010), or is at least 95% identical to (S), where a change in
XX expression is indicative of anti-neoplastic activity. (I) has cytostatic
XX activity and can be used in gene therapy. M1 can be used for screening
XX an anti-neoplastic agent, and can be used for producing a product which
XX is the data collected with respect to the anti-neoplastic agent as a
XX result of M1, and the data is sufficient to convey the chemical
XX structure and/or properties of the agent. M1 can be used in the
XX treatment of cancer such as colon, breast, stomach, lung, thyroid,
XX oesophageal, ovarian, kidney, prostate or pancreatic cancer,
XX adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
XX infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
XX carcinoma, papillary carcinoma and Wilm's tumour.
XX
XX SQ Sequence 3454 BP; 594 A; 1217 C; 1136 G; 507 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 2,79e-34 Length: 3454
XX Score: 392.50 Matches: 88
XX Percent Similarity: 56.68% Conservative: 35
XX Best Local Similarity: 40.55% Mismatches: 77
XX Query Match: 31.35% Indels: 17
XX DB: 24 Gaps: 6
XX
XX US-09-830-144-2_COPY_76_303 (1-228) x ABL70018 (1-3454)
Qy 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCys 21
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 721 GAAGCCCGGCTCTTTGGAGCCCTGCGACCCCAACATAATTGCGCTTAGGGGCGCTGC 780
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 22 LeuAsnPro-----ValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsn 39
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 781 CTCAACCCCCACACCTCTGCTAGTGATGGAGATATGCCCGGGGTGGTGGCTAGGAGCAGG 840
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 40 ValLeuHisGlyAlaGluProLeuProTyrThrAlaAlaHisAlaMetSerTrpCys 59
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 841 GTGCTGGCAGGTGCGCGGGTGCCACCTCACGTG-----CTGTCACTGGGCT 888
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 60 LeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHis 79
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
Db      889  GTCCAGGTGCCCCGGGCGATGAACCTACCTACACAAATGATGCCCTCTGTGCCCATCATCCAC 948
Qy      80   ArgAspLeuLysProProAsnLeuLeuVal-----AlaGly 92
Db      949  CCGGACCTCAAGTCCATCAACATCTGATCGTCGAGGCCATCGAGAACCAACCTCGCA 1008
Qy      93   GlyThrValLeuLysIleCysAspPheGlyThrAlaCysAsp---IleGlnThrHisMet 111
Db      1009 GACACGGTGTCAAGATCAACGCGACTTCGGCTCGCCCGGAGTGGCACAAGACCAACCAAG 1068
Qy      112  ThrAspAsnLysGlySerAlaAlaATPmetAlaProGluValPheGluLysSerAsnTyr 131
Db      1069 ATGAGCGCTCGCGGAGACCTACACGCTGATGCGCCCGGAGATTATCGTCTCTCCCTTTC 1128
Qy      132  SerGluLysCysAspValPheSerTyrGlyIleIleLeuTyrGluValIleThrArgArg 151
Db      1129  TCCAAAAGCAGTGTATGTGAGCTTCGGGGTGCTGTGGAGAGCTGTACGGGGGAG 1188
Qy      152  LysProPheAspGluIleGlyGlyProAlaPheArgIleMetTyrAlaVal---HisAsn 170
Db      1189 GTCCCTTACCGTAGATC-----GAGGCTTGGCCGCTGATGGCGGTATGAT 1242
Qy      171  GlyThrArgProProLeuIleLysAsnLeuProLysProIleGluSerLeuMetThrArg 190
Db      1243 AAGCTAGCGCTGCCCATTCCTCCACAGTGCCTCCGAGCCTTTGCCCGCTCTCGAGGAA 1302
Qy      191  CysTyrSerLysAspProSerGlnArgProSerMetGluIleValIys 207
Db      1303 TCTGTGGACCCAGACCCCGGCGGCGCAGATTTCGTAGCATCTTGAG 1353

RESULT 14
AAC43254
ID      AAC43254 standard: DNA; 1224 BP.
XX
XX      AAC43254;
XX
XX      17-OCT-2000 (first entry)
XX
DE      Arabidopsis thaliana DNA fragment SEQ ID NO: 38594.
XX
XX
XX      Hybridisation assay; genetic mapping; gene expression control;
XX      protein identification; signal transduction pathway;
XX      metabolic pathway; promoter; termination sequence; ss.
OS      Arabidopsis thaliana.
XX
XX      EP1033405-A2.
XX
XX      06-SEP-2000.
XX
XX      25-FEB-2000; 2000EP-0301439.
XX
XX      25-FEB-1999; 99US-0121825.
XX      05-MAR-1999; 99US-0123180.
XX      09-MAR-1999; 99US-0123548.
XX      23-MAR-1999; 99US-0125788.
XX      25-MAR-1999; 99US-0126264.
XX      29-MAR-1999; 99US-0126785.
XX      01-APR-1999; 99US-0127462.
XX      06-APR-1999; 99US-0128234.
XX      08-APR-1999; 99US-0128714.
XX      16-APR-1999; 99US-0129845.
XX      19-APR-1999; 99US-0130077.
XX      21-APR-1999; 99US-0130449.
XX      23-APR-1999; 99US-0130510.
XX      23-APR-1999; 99US-0130891.
XX      28-APR-1999; 99US-0131449.
XX      30-APR-1999; 99US-0132048.
XX      30-APR-1999; 99US-0132407.
XX      04-MAY-1999; 99US-0132484.
XX      05-MAY-1999; 99US-0132485.
XX      06-MAY-1999; 99US-0132486.
```

```
PR      06-MAY-1999; 99US-0132487.
PR      07-MAY-1999; 99US-0132863.
PR      11-MAY-1999; 99US-0134256.
PR      14-MAY-1999; 99US-0134218.
PR      14-MAY-1999; 99US-0134219.
PR      14-MAY-1999; 99US-0134221.
PR      14-MAY-1999; 99US-0134370.
PR      18-MAY-1999; 99US-0134768.
PR      19-MAY-1999; 99US-0134941.
PR      20-MAY-1999; 99US-0135124.
PR      21-MAY-1999; 99US-0135324.
PR      24-MAY-1999; 99US-0135629.
PR      25-MAY-1999; 99US-0136021.
PR      27-MAY-1999; 99US-0136392.
PR      28-MAY-1999; 99US-0136782.
PR      01-JUN-1999; 99US-0137222.
PR      03-JUN-1999; 99US-0137528.
PR      04-JUN-1999; 99US-0137502.
PR      07-JUN-1999; 99US-0137724.
PR      08-JUN-1999; 99US-0138094.
PR      10-JUN-1999; 99US-0138540.
PR      10-JUN-1999; 99US-0138847.
PR      14-JUN-1999; 99US-0139119.
PR      16-JUN-1999; 99US-0139452.
PR      17-JUN-1999; 99US-0139453.
PR      17-JUN-1999; 99US-0139492.
PR      18-JUN-1999; 99US-0139454.
PR      18-JUN-1999; 99US-0139455.
PR      18-JUN-1999; 99US-0139456.
PR      18-JUN-1999; 99US-0139457.
PR      18-JUN-1999; 99US-0139458.
PR      18-JUN-1999; 99US-0139459.
PR      18-JUN-1999; 99US-0139460.
PR      18-JUN-1999; 99US-0139461.
PR      18-JUN-1999; 99US-0139462.
PR      18-JUN-1999; 99US-0139463.
PR      18-JUN-1999; 99US-0139750.
PR      18-JUN-1999; 99US-0139763.
PR      21-JUN-1999; 99US-0139817.
PR      22-JUN-1999; 99US-0139899.
PR      23-JUN-1999; 99US-0140353.
PR      23-JUN-1999; 99US-0140354.
PR      24-JUN-1999; 99US-0140695.
PR      28-JUN-1999; 99US-0140823.
PR      29-JUN-1999; 99US-0140991.
PR      30-JUN-1999; 99US-0141287.
PR      01-JUL-1999; 99US-0141842.
PR      01-JUL-1999; 99US-0142154.
PR      02-JUL-1999; 99US-0142055.
PR      06-JUL-1999; 99US-0142390.
PR      08-JUL-1999; 99US-0142803.
PR      09-JUL-1999; 99US-0142920.
PR      12-JUL-1999; 99US-0142977.
PR      13-JUL-1999; 99US-0143542.
PR      14-JUL-1999; 99US-0143624.
PR      15-JUL-1999; 99US-0144005.
PR      16-JUL-1999; 99US-0144085.
PR      16-JUL-1999; 99US-0144086.
PR      19-JUL-1999; 99US-0144325.
PR      19-JUL-1999; 99US-0144331.
PR      19-JUL-1999; 99US-0144332.
PR      19-JUL-1999; 99US-0144333.
PR      19-JUL-1999; 99US-0144334.
PR      19-JUL-1999; 99US-0144335.
PR      20-JUL-1999; 99US-0144352.
PR      20-JUL-1999; 99US-0144632.
PR      21-JUL-1999; 99US-0144884.
PR      21-JUL-1999; 99US-0144814.
PR      21-JUL-1999; 99US-0145086.
PR      21-JUL-1999; 99US-0145088.
PR      22-JUL-1999; 99US-0145085.
PR      22-JUL-1999; 99US-0145087.
PR      22-JUL-1999; 99US-0145089.
```


AC AAC39537;
XX
XX 17-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 24996.
DE
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131445.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138099.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139451.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145813.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.


```
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

Alignment Scores:

```
Pred. No.: 1.64e-32
Score: 373.00
Percent Similarity: 57.80%
Best Local Similarity: 39.91%
Query Match: 29.79%
DB: 21
```

US-09-830-144-2_COPY_76_303 (1-228) x AAC39537 (1-1631)

```
QY 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCys 21
Db 773 GAAGTTTCTATGCTTGATCTTTGAAGCATCTTAAACATCGTTAGGTTTATTGGTCGGTGC 832
QY 22 LeuAsnProVal-----CysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsn 39
Db 833 ATTAACCGATGCTGGTGCATCGTGACTGAATATGCAAAAGGAGGGTCTGTGCACACAG 892
QY 40 ValLeu-----HisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAla 55
```

```
Db 893 TTTCTGACTAAGAGACAAAACCGAGCTGTCCTTTTGAAGTTA-----GCT 937
QY 56 MetSerTrpCysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLys 75
Db 938 GTTATGACAGCGGTTGGATGTTGCCAGGGGTATGCTTACGTCATGAG-----CGC 988
QY 76 AlaLeuIleHisArgAspLeuLysProAsnLeuLeuValAlaGlyGlyThrVal 95
Db 989 AACTTTATACACCGGGATCTAAAGTCAGATAACCTCTCATATCAGCTGATCGTCCATC 1048
QY 96 LeuLysIleCysAspPheGlyThrAla---CysAspIleGlnThrHis---MetThrAsn 113
Db 1049 ---AAGATTGCTGATTTTGGTGTGCAAGATTGAAGTTCAAACCGAAGGATGACACCA 1105
QY 114 AsnLysGlySerAlaAlaTrpMetAlaProGluValPheGluGlySerAsnTyrSerGlu 133
Db 1106 GAGACTGGAACTTACAGATGGATGGCACAGAGATGATCCAGCACACACCTTACACTCAA 1165
QY 134 LysCysAspValPheSerTrpGlyIleIleLeuTrpGluValIleThrArgLysPro 153
Db 1166 AAGTGGACGTGTATAGTTTGGAAATCGTCTGTGGAGTTGATTACAGGTCTGTTACCG 1225
QY 154 PheAspGluIleGlyGly-----ProAlaPheArgIleMetTrpAlaValHisAsnGly 171
Db 1226 TTCAGAAATGACGGCGGTTTCAGGTGCATTTGCAGTG-----GTCAACAGAGGA 1276
QY 172 ThrArgProProLeuIleLysAsnLeuProLysProIleGluSerLeuMetThrArgCys 191
Db 1277 GTCCGTCCACAGTCCCAGCAGATGTCCTCTGTGCTTGGAGAGATCATGACACGTTC 1336
QY 192 TrpSerLysAspProSerGlnArgProSerMetGluIleValLysIleMet 209
Db 1337 TGGATGCGGACCCCTGAAGTCGCTTGTGTTTGCAGAGATTGTCAATCTTCTG 1390
```

Search completed: December 10, 2002, 04:05:33

Job time : 328 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 10, 2002, 03:59:01 ; Search time 2220 Seconds
(without alignments)
1663.320 Million cell updates/sec
Title: US-09-830-144-2_COPY_76_303
Perfect score: 1252
Sequence: 1 VELRQLSRVNHNPVIVKLYG.....MTHLMRYFPGADEPLQYPCQ 228

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlp
-MODE=frame_p2n.model -DEV=xlp
-Q/cgn2_1/USPTO.spool/US09830144/runat_04122002_141754_4354/app_query.fasta_1.391
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09830144 @CNG_1_1_1716 @runat_04122002_141754_4354 -ICPU=3
-NO_XLPXY -NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:.*
1: em_estba:.*
2: em_esthum:.*
3: em_estin:.*
4: em_estmu:.*
5: em_estov:.*
6: em_estpl:.*
7: em_estro:.*
8: em_htc:.*
9: gb_est1:.*
10: gb_est2:.*
11: gb_htc:.*
12: gb_est3:.*
13: gb_est4:.*
14: gb_est5:.*
15: em_estfun:.*
16: em_estom:.*
17: gb_gss:.*
18: em_gss_hum:.*
19: em_gss_inv:.*
20: em_gss_pln:.*
21: em_gss_vrt:.*
22: em_gss_fun:.*
23: em_gss_mam:.*
24: em_gss_mus:.*
25: em_gss_other:.*
26: em_gss_pro:.*
27: em_gss_rod:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1167	93.2	739	13	BI696710	BI696710 603347214
2	1069	85.4	604	13	BM426610	BM426610 pfgfn.pk0
3	1016	81.2	1062	13	BM554120	BM554120 AGENCOURT
4	984	78.6	646	13	BJ073883	BJ073883 BJ073883
5	982	78.4	688	13	BJ062988	BJ062988 BJ062988
6	982	78.4	696	13	BJ074867	BJ074867 BJ074867
7	953.5	76.2	910	12	BF780358	BF780358 602103276
8	932.5	74.5	998	9	AL525728	AL525728 AL525728
9	847	67.7	542	10	AW960377	AW960377 EST372448
10	823	65.7	641	14	BQ386875	BQ386875 NISC_mn20
11	787	62.9	827	9	AL550589	AL550589 AL550589
12	753	60.1	1006	14	BQ219348	BQ219348 AGENCOURT
13	748	59.7	686	12	BG548917	BG548917 602575750
14	741	59.2	771	13	BI093821	BI093821 602860601
15	737	58.9	682	10	AV398933	AV398933 AV398933
16	715	57.1	586	13	BJ035129	BJ035129 BJ035129
17	707	56.5	681	10	AV398935	AV398935 AV398935
18	705	56.3	929	9	AL520975	AL520975 AL520975
19	631.5	50.4	1054	14	BQ067693	BQ067693 AGENCOURT
20	631	50.4	607	9	AI541706	AI541706 SD06739.5
21	628	50.2	594	13	BJ014293	BJ014293 BJ014293
22	587	46.9	747	13	BI334774	BI334774 602999677
23	559	44.6	516	13	BJ006787	BJ006787 BJ006787
24	557	44.5	858	13	BG916821	BG916821 602816079
25	553	44.2	711	10	AV718168	AV718168 AV718168
26	549.5	43.9	760	12	BG773488	BG773488 602720209
27	549	43.8	822	13	BI757369	BI757369 603029591
28	548	43.8	538	14	BQ557823	BQ557823 H4048E02-
29	540	43.1	745	13	BI767913	BI767913 603061057
30	515.5	41.2	647	13	BJ008789	BJ008789 BJ008789
31	497	39.7	599	10	BE284386	BE284386 601087550
32	496	39.6	286	9	AA450856	AA450856 VG55H03.Y
33	483	38.6	585	9	AL773651	AL773651 AL773651
34	474	37.9	718	9	AL652289	AL652289 AL652289
35	469.5	37.5	1191	13	BM472107	BM472107 AGENCOURT
36	457	36.5	1033	12	BE882182	BE882182 601505283
37	452	36.1	647	10	BE665645	BE665645 BB665645
38	445	35.5	569	9	AU177912	AU177912 AU177912
39	443	35.4	338	12	BE707048	BE707048 MRO-HT040
40	438	35.0	831	14	BQ736206	BQ736206 AGENCOURT
41	437.5	34.9	845	12	BF027003	BF027003 601671216
42	430	34.3	941	12	BG029948	BG029948 602297101
43	420	33.5	248	9	AI614762	AI614762 VG55H03.Y
44	387.5	31.0	689	14	BQ745060	BQ745060 UI-M-EHOP
c 45	374	29.9	740	13	BJ441629	BJ441629 BJ441629

ALIGNMENTS

RESULT 1
BI696710
LOCUS 603347214F1 NCI_CGAP_Mam2 Mus musculus cDNA clone IMAGE:5374971 5',
DEFINITION 739 bp mRNA linear EST 18-SEP-2001
mRNA sequence.
ACCESSION BI696710
VERSION BI696710.1 GI:15659339
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 739)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
DNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: LLM11954 row: k column: 04
High quality sequence stop: 717.
Location/Qualifiers
1. 739
/organism="Mus musculus"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone="IMAGE:5374971"
/clone_lib="NCI CGAP Mam2"
/tissue_type="tumor, biopsy sample"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: PCMV-SPORE6; Site: 1: Sal1;
Site 2: NotI; Cloned unidirectionally. Primer: oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

BASE COUNT 195 a 169 c 188 g 187 t

ORIGIN

Alignment Scores:
Pred. No.: 1,62e-124 Length: 739
Score: 1167.00 Matches: 220
Percent Similarity: 98.65% Conservative: 0
Best Local Similarity: 98.65% Mismatches: 1
Query Match: 93.21% Indels: 3
DB: 13 Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x B1696710 (1-739)

QY 8 ArgValaAenHispRoasnlleVallyseuTyrgYlaAaCysLeuAsnProvalCysleu 27
Db 3 CGTGTAAACCATCTTAACATGTGTCAAGTTGTACGAGCCTGCTGAATCCAGTATGTCTT 62
QY 28 ValMetGluTyAlaGluGlyGlySerLeuTyrgAsnValleuHisGlyAlaGluProleu 47
Db 63 GGTATGAATATGACAGAGGGGGCTCATTTGTAATATGCTGATGATGCTGAACCATTTG 122
QY 48 ProTyTyTyThAlaAlaHisAlaMetSerTrpCysLeuGlnCysSerGlnGly-ValAl 67
Db 123 CCTTACTACACTGCTGCTATCCCATGAGCTGCTTTTACAGTGTCCCAAGTATGTGC 182
QY 67 aTyrlleuHisSerMetGlnProlyAlaLeuIleHisArgAspLeuLysProProAsnle 87
Db 183 TTACCTGCACACATGACGCCCAAGCCGTGATTCACAGGGACCTCAAGCTCCAACTT 242
QY 87 lleuLeuValAlaGlyGlyThValleuLyslleCysAspPheGlyThrAlaCysAspI 107
Db 243 GCTGCTGGTGGAGAGGACAGTCTTAATAATCTGCATTTTGTATACAGCTTGACAT 302
QY 107 eGlnThrHisMetThrAsnLysGlySerAlaAlaIleTrpMetAlaProGluValAlPheG 127
Db 303 CCAACACACATGACCATTAATAAGGAGTCTGCTTGATGGGCTCGAAGTCTTTGA 362
QY 127 uGlySerAsnTyrgSerGluLysCysAspValPheSerTrpGlyllelleuTyrgLys 147
Db 363 AGGTGCAATTAAGTGAAGAGTGTGATGCTTCAGCTGGGGTATTATCTCTGGGAAGT 422
QY 147 lleThrArgArgLysProPheAspGluilleGlyGlyProAlaPheArgIleMetTrpAl 167
Db 423 GATTAACAGCCCGAATCCCTTGATGATGAGATCGGTGGCCAGCTTCAAAATCATGTGGC 482
QY 167 aValHisAsnGlyThArgProProleuIleLysAsnLeuProLysProIleGluSerle 187

Db 483 TGTTCATATGACACTGCACCACTGATCAAAATTTTACTAAGCCCATTTGAGACTT 542
QY 187 uMetThrArgCys-TrpSerLysAspProSerGlnArgProSerMetGluGluIleVal 207
Db 543 GATGACACGCTGTGTGTGTCTTAAGACCAATCTCAGGC-CCTTCAATGGAGAAATGTGA 601
QY 207 yslleMetThrHisLeuMetArgTyrgPheProGlyAlaAspGluProleuGlnTyrgProc 227
Db 602 AAATTAATGACTCTGATGCGGTACTTCCAGAGCGATGAGCCATTAAGTATCTCT 661
QY 227 ysgln 228
Db 662 GTCAAG 666

RESULT 2
BM426610
LOCUS
DEFINITION
BM426610
pgf2n.pk004.d15 Normalized Chicken Abdominal Fat library (pgf2n)
Gallus gallus CDNA clone pgf2n.pk004.d15 5' similar to
emb|CAB87605.1 (A1121964) d015461.1.4 (mitogen-activated protein
kinase kinase kinase 7 (TGF-beta activated kinase 1d (TAK1))) [Homo
sapiens], mRNA sequence.
BM426610
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
chicken.
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 604)
Cogburn,L.A., Morgan,R. and Burnside,J.
ESTs from Normalized Chicken fat CDNA library-USDA/IFAFS Animal
Genome Project
Unpublished (2002)
Contact: Larry A. Cogburn
University of Delaware
Townsend Hall, Newark, DE 19717, USA
Tel: 302-831-1335
Fax: 302-831-2822
Email: cogburnudel.edu, www.chickest.udel.edu.

FEATURES
source
1. 604
/organism="Gallus gallus"
/strain="Commercial broiler, Ottawa Research Centre,
leghorn"
/db_xref="taxon:9031"
/clone="pgf2n.pk004.d15"
/clone_lib="Normalized Chicken Abdominal Fat library
(pgf2n)"
/sex="Male and Female"
/tissue_type="Abdominal Fat"
/dev_stage="Embryonic (d18,d19); post-hatch (d1,w3,w7,w9
w16,lyr)"
/lab_host="E. coli EMDH10B"
/note="Vector: PCMVSPORT6; Library made from equivalent
pools of total RNA isolated from each developmental age
(across strains); Single pass sequencing from 5'-end"

BASE COUNT 167 a 125 c 140 g 167 t 5 others

ORIGIN

Alignment Scores:
Pred. No.: 2,6e-113 Length: 604
Score: 1069.00 Matches: 194
Percent Similarity: 98.99% Conservative: 2
Best Local Similarity: 97.98% Mismatches: 2
Query Match: 85.38% Indels: 0
DB: 13 Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x BM426610 (1-604)

QY 30 GluTyAlaGluGlyGlySerLeuTyrgAsnValleuHisGlyAlaGluProleuProTyrg 49
|||||

Db 2 GAGTATGCTGAGGAGTTCTCTGTACAAATGCTGTGCAATGCTGCTGAACCTCTCTCCCTCAT 61
 Qy 50 TTTTThrAlaAlaHisAlaMetSerTrpCysLeuGlnCysSerGlnGlyValAlaTyrLeu 69
 Db 62 TATATCTGCTGCACCGCAATGAGTTGGTTTACAGTGTTCCTCCAGAGTGGCAATCTT 121
 Qy 70 HisSerMetGlnProLysAlaAlaLeuHisArgAspLeuLysProProAsnLeuLeu 89
 Db 122 CACAGTATGAACCAACCAAGCCCTAATTCACAGAGACCTGAAACACCAAAATTCCTCTTG 181
 Qy 90 ValAlaGlyGlyThrValLeuLysLeuCysAspPheGlyThrAlaCysAspIleGlnThr 109
 Db 182 GTAGCTGGGGGACAGTTCTTAAAGATCTGTGATTTTGTTCAGCTGTGATATTCACAA 241
 Qy 110 HisMetThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGluValPheGluGlySer 129
 Db 242 CACATGACCAACAAATPAGGGAAGTCTGCTGGATGGACCTGAAGTTTGTGAGGTAGC 301
 Qy 130 AsnTyrSerGluLysCysAspValPheSerTrpGlyIleLeuLeuTrpGluValIleThr 149
 Db 302 AATTACAGTGAACCAATGTACGCTTTTTCAGTTGGGGTATTAATCTTTGGGAGGTAATCACC 361
 Qy 150 ArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMetTrpAlaValHis 169
 Db 362 CGTAGGAAACCTTTTTCATGAGATTTGGTCCAGCTTTCCGCAATTAATGTGGCGATTCAC 421
 Qy 170 AsnGlyThrArgProProLeuLysAsnLeuProLysProIleGluSerLeuMetThr 189
 Db 422 AATGTACTCGACCACTGATCAAAACTTACCTAAACCAATTCAGAGTTAATGACC 481
 Qy 190 ArgCysTrpSerLysAspProSerGlnArgProSerMetGluGluIleValLysIleMet 209
 Db 482 CGCTGTGGTCAAGGATCCCTCACACGACCTTCATGAGGAGAAATGTTAAAAATAATG 541
 Qy 210 ThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyrProCys 227
 Db 542 ACACACTTTCATCGGTACTTTCCNNNAGCTGATGAACCTCTCGCATATCCCTTGC 595

RESULT 3
 BM554120 1062 bp mRNA linear EST 20-FEB-2002
 LOCUS AGENCOURT_6546903 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5742353
 5', mRNA sequence.

ACCESSION BM554120
 VERSION BM554120.1 GI:18793446
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1062)
 AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLNL2760 row: f column: 18
 High quality sequence stop: 730.

FEATURES
 Location/Qualifiers
 1..1062
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5742353"
 /clone_lib="NIH_MGC_119"
 /tissue_type="medulla"
 /lab_host="DH10B"

SOURCE

RESULT 4

BJ073883

LOCUS

DEFINITION

ACCSSION

VERSION

BJ073883 646 bp mRNA linear EST 11-DEC-2001
 BJ073883 NIBB Mochii normalized xenopus tailbud library Xenopus
 laevis cDNA clone XL101n05 5', mRNA sequence.
 BJ073883
 BJ073883.1 GI:17504072

/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: EcoRV (destroyed); RNA source normal medulla from
 anonymous male age 27. Library is oligo-dr primed and
 directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.3 kb, insert size range
 0.9-3 Kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 013. Note:
 this is a NIH MGC Library."

BASE COUNT 253 a 243 c 298 g 266 t 2 others
 ORIGIN

Alignment Scores:

Pred. No.: 8.56e-107 Length: 1062
 Score: 1016.00 Matches: 203
 Percent Similarity: 94.52% Conservative: 4
 Best Local Similarity: 92.69% Mismatches: 5
 Query Match: 81.15% Indels: 7
 DB: 13 Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x BM554120 (1-1062)

Qy 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20

Db 408 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCATCTCTAATATTGTAAGCTTTATGAGGCC 467

Qy 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnVal 40

Db 468 TGCTTGAATCCAGTGTCTTTGTGATGGAATATGCTGAAGGGGCTCTTTATATAATGTG 527

Qy 41 LeuHisGlyAlaGluProLeuProTyrThrAlaAlaHisAlaMetSerTrpCysLeu 60

Db 528 CTGCATGGTCTGAAACCATTTGCCATATTATCTGCTGCTGCCACGCAATGAGTTGGTGT 587

Qy 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80

Db 588 CAGTGTTCCTCCAGGAGTGGCTTATCTTCACAGCATGCAACCCCAAGCGCTAATTCACAGG 647

Qy 81 AspLeuLysProProAsnLeuLeuValAlaGlyThrValLeuLysIleCysAsp 100

Db 648 GACCTGAACACCACTTACTGCTGTTGTCAGGGGGGACAGTCTCTAAAAATTTGTGAT 707

Qy 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 120

Db 708 TTTGTTACAGCTGTGACATTCAGACACATGACCAATAACCAAGGGGAGTGTCTGTGG 767

Qy 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 140

Db 768 ATGCACCTGAAGTTTTTGAAGGTAGTAATTACAGTGAATAAATGTGACGCTTCACGCTGG 827

Qy 141 GlyIleLeuTrpGluValIleThrArgArg-LysProPheAspGluIleGly-P 160

Db 828 GGTATTATTCTTTGGGAAGTGATAACGCGTCGGAAAAACCCCTTTGATGAGATGGGTGGCCC 887

Qy 160 roAlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeu-IleLysAsn 179

Db 888 CAGCTTTCCGAATCATGTGGCTGGTCATATGCTACTCGACCACTGATATAAAAAAT 947

Qy 180 LeuProLysProIleGlu-SerLeuMetThr-ArgCysTrpSerLysAspPro-SerGln 198

Db 948 TTACTAGGCCCATTAAGAACCTGATGATCCCTGTTGGTGGTAAAGATCTTTCCAG 1007

Qy 199 ArgProSerMetGlu-GluIleValLysIleMetThrHisLeu 212

Db 1008 CGCCCTTTCATGGAGGGAATGGTGAATAAATGACCTCACTTG 1050

KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
Xenopodinae; Xenopus.
1 (bases 1 to 646)
REFERENCE Kiteyama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara
AUTHORS Y.
TITLE Expressed genes in X. laevis embryo
JOURNAL Unpublished (2001)
COMMENT Contact: Tadaeu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
source
1..646
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone_lib="X101n05"
/clone_lib="NIBB Mochii normalized Xenopus tailbud
library"
/tissue_type="whole embryo"
/dev_stage="stage 25"
/note="vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; cDNAs
were oligo-dT primed and directionally cloned. Staging
according to Nieuwkoop and Faber. Library is substracted
and was constructed by N. Garrett and A.M. Zorn,
(Wellcome/CRC Institute). "

BASE COUNT 190 a 132 c 143 g 181 t
ORIGIN

Alignment Scores:
Pred. No.: 1,99e-103 Length: 646
Score: 984.00 Matches: 182
Percent Similarity: 97.37% Conservative: 3
Best Local Similarity: 95.79% Mismatches: 5
Query Match: 78.59% Indels: 1
DB: 13 Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x BU073883 (1-646)

QY 39 AenVallleuHIGLYAlaGluProLeuProTyTYrThraAlaAlaHisAlaMetSerTrp 58
Db 4 AATGTTTGCATGAGAGCTGAACCTTGCTTACAT-ACGTGCTCCCATGATGATTGG 62
QY 59 CysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIle 78
Db 63 TGTTTACAAATGTCGCAAGAGGATTCATATTTACATACATGAAGCCAAAGGCTCTGATT 122
QY 79 HisArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIle 98
Db 123 CACAGGAGCCTCAAAACCAAACTTGTTCGGTGAAGTGGAGGAGCAGCTTCTTAAGATT 182
QY 99 CysAspPheGlyThrAlaCysAspIleGlnThrHisMetCThrAsnAsnLysGlySerAla 118
Db 183 TGTGACTTTGGTACAGCTGTGATTCAGACTCAGACTCAATTAACAAAGAAAGTGCA 242
QY 119 AlaTrpMetAlaProGluValPheGluGlySerAsnTyrSerGlnLysCysAspValPhe 138
Db 243 GCATGAGTGGCTCCAGAACTTTTGAAGTAGCACTACAGCAAAATAATGTGACGTGTTT 302
QY 139 SerTrpGlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGly 158
Db 303 AGTTGGGCAATTATCTTTGGGAAGTATATAACCGAATAAACTTTGAGAAATTGGT 362
QY 159 GlyProAlaPheArgIleMetCThrAlaValHisAsnGlyThrArgProProLeuIleLys 178
Db 363 GGTCCAGGCTTCGTATATATGTGGGCTGTTCCAAATGCTACTCGGCACCATTAATTTAAA 422

QY 179 AsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGln 198
Db 423 AATTTGGCTTAAGCCTATTGAAAGCTTATGATGACTGCTGCTGCCAAGATCCCCACAA 482
QY 199 ArgProSerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGly 218
Db 483 AGACCTTCAAATGAGAGGATTTGTCAGATATATGACATCTTAAGCAGATATTTCTTGGA 542
QY 219 AlaAspGluProLeuGlnTyrProCysGln 228
Db 543 GCAGACGTTTCCTTACAGTATCCTTTGTCAG 572

RESULT 5
LOCUS BU062988 688 bp mRNA linear EST 10-DEC-2001
DEFINITION BU062988 NIBB Mochii normalized Xenopus tailbud library Xenopus
laevis cDNA clone X1069m17 5', mRNA sequence.
ACCESSION BU062988
VERSION BU062988.1 GI:17470746
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
Xenopodinae; Xenopus.
1 (bases 1 to 688)
REFERENCE Kiteyama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara
AUTHORS Y.
TITLE Expressed genes in X. laevis embryo
JOURNAL Unpublished (2001)
COMMENT Contact: Tadaeu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
source
1..688
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone_lib="X1069m17"
/clone_lib="NIBB Mochii normalized Xenopus tailbud
library"
/tissue_type="whole embryo"
/dev_stage="stage 25"
/note="vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; cDNAs
were oligo-dT primed and directionally cloned. Staging
according to Nieuwkoop and Faber. Library is substracted
and was constructed by N. Garrett and A.M. Zorn,
(Wellcome/CRC Institute). "

BASE COUNT 206 a 139 c 154 g 188 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 3,74e-103 Length: 688
Score: 982.00 Matches: 177
Percent Similarity: 97.83% Conservative: 3
Best Local Similarity: 96.20% Mismatches: 4
Query Match: 78.43% Indels: 0
DB: 13 Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x BU062988 (1-688)

QY 45 GluProLeuProTyTYrThraAlaAlaHisAlaMetSerTrpCysLeuGlnCysSerGln 64
Db 22 GAACCTTTCCTTACATATCTGCTGCCATGCAAGAGTGGTGTTTTCAATGTCCCA 81
QY 65 GlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysPro 84
Db 82 GGAGTTGCATATTTACATGATGAAGCCAAAGGCTGTGATTCACAGGAGCCTCAACCA 141
QY 85 ProAsnLeuLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAla 104

```

Db 142 CCAAACTTGTGTGCTAGCTGGAGGCACTGTTCTTAAGATTGTGACTTTGGTACAGCC 201
Qy 105 CysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGlu 124
Db 202 TGTGATATTACAGACTCACATGACTAATAACAAGGAAGTGCAGCATGGATGCTCCAGAA 261
Qy 125 ValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrpGlyIleIleLeu 144
Db 262 GTTTTGAAGGTAGCAACTACAGCGAATAATGTGACGTGTTAGTTGGGCATTATTCTT 321
Qy 145 TrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIle 164
Db 322 TGGGAAGTAATAACCCGAAGAAACCTTTCGATGAAATGGTGGTCCAGCGTTCGGTATA 381
Qy 165 MetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIle 184
Db 382 ATGTGGGCTGTTACAAATGGTACTCGGCCACCATTAATTAATAAATTTGCTTAAGGCTATT 441
Qy 185 GluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgProSerMetGluGlu 204
Db 442 GAAAGCTTAATGACTCGCTGCTGGTCCAAAGATCCCCCAAGACCTTCAATGGAGGAG 501
Qy 205 IleValIleMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGln 224
Db 502 ATTGTCAAGATAATGACATCTAAAGCAGTATTTTCTCGGAGCAGACGTTTCTTACAG 561
Qy 225 TyrProCysGln 228
Db 562 TATCCTTGTCAG 573

RESULT 6
BJ074867
LOCUS BJ074867
DEFINITION BJ074867 NTBB Mochii normalized Xenopus tailbud library Xenopus laevis cDNA clone XL071110 5', mRNA sequence.
ACCESSION BJ074867
VERSION BJ074867.1 GI:17505056
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae; Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 696)
AUTHORS Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara,Y.
TITLE Expressed genes in X. laevis embryo
JOURNAL Unpublished (2001)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
FEATURES
Location/Qualifiers
1..696
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="XL071110"
/clone_lib="NIBB Mochii normalized Xenopus tailbud library"
/tissue_type="whole embryo"
/vec_stage="stage 25"
/notes="Vector: pBSRN3; Site 1: NotI; Site 2: EcoRI; cDNAs were oligo-dT primed and directionally cloned. Staging according to Nieuwkoop and Faber. Library is substracted and was constructed by N. Garrett and A.M. Zorn, (Wellcome/CRC Institute)."
BASE COUNT 210 a 142 c 155 g 188 t 1 others
ORIGIN

```

```

Alignment Scores:
Pred. No.: 3..81e-103 Length: 696
Score: 982.00 Matches: 177
Percent Similarity: 97.83% Conservative: 3
Best Local Similarity: 96.20% Mismatches: 4
Query Match: 78.43% Indels: 0
DB: 13 Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x BJ074867 (1-696)

Qy 45 GluProLeuProTyrTrpThrAlaAlaHisAlaMetSerTrpCysLeuGlnCysSerGln 64
Db 22 GAACCTTTTGCTTACTACTGCTGCCATGCAATGAGTTGGTGTTTACAAATGTGCCAA 81
Qy 65 GlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysPro 84
Db 82 GGAGTTGCATATTTACATAGCATGAAGCCAAAGGCTCTGATTCACAGGGACCTCAACCA 141
Qy 85 ProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAla 104
Db 142 CCAAACTTGTGTGCTAGCTGGAGGCACTGTTCTTAAGATTGTGACTTTGGTACAGCC 201
Qy 105 CysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGlu 124
Db 202 TGTGATATTACAGACTCACATGACTAATAACAAGGAAGTGCAGCATGGATGCTCCAGAA 261
Qy 125 ValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrpGlyIleIleLeu 144
Db 262 GTTTTGAAGGTAGCAACTACAGCGAATAATGTGACGTGTTAGTTGGGCATTATTCTT 321
Qy 145 TrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIle 164
Db 322 TGGGAAGTAATAACCCGAAGAAACCTTTCGATGAAATGGTGGTCCAGCGTTCGGTATA 381
Qy 165 MetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIle 184
Db 382 ATGTGGGCTGTTCAAAATGGTACTCGGCCACCATTAATTAATAAATTTGCTTAAGGCTATT 441
Qy 185 GluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgProSerMetGluGlu 204
Db 442 GAAAGCTTAATGACTCGCTGCTGGTCCAAAGATCCCCCAAGACCTTCAATGGAGGAG 501
Qy 205 IleValIleMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGln 224
Db 502 ATTGTCAAGATAATGACATCTAAAGCAGTATTTTCTCGGAGCAGACGTTTCTTACAG 561
Qy 225 TyrProCysGln 228
Db 562 TATCCTTGTCAG 573

RESULT 7
BJ0780358
LOCUS BJ0780358
DEFINITION 602103276F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:4221379 5', mRNA sequence.
ACCESSION BF780358
VERSION BF780358.1 GI:12085481
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 910)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Jeffrey B. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
Clone distribution: MGC clone distribution information can be

```

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov
plate: LLM9806 row: h column: 20
High quality sequence stop: 664.
Location/Qualifiers

FEATURES

source

1..910

/organism="Mus musculus"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone_image="4221379"

/clone_lib="NCI CGAP Kid4"

/lab_host="DH10B (T1 phage-resistant)"

/note="Organ: kidney; Vector: pCMV-SPORT6; Site 1: NotI;

Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.

Average insert size 1.75 kb. Constructed by Life

Technologies. Note: this is a NCI CGAP Library. |"

BASE COUNT 247 a 209 c 246 g 207 t 1 others

ORIGIN

Alignment Scores:

Pred. No.:	1.15e-99	Length:	910
Score:	953.50	Matches:	194
Percent Similarity:	92.86%	Conservative:	1
Best Local Similarity:	92.38%	Mismatches:	6
Query Match:	76.16%	Indels:	12
DB:	12	Gaps:	1

US-09-830-144-2_COPY_76_303 (1-228) x BF780358 (1-910)

```
OY 28 ValmetGluTyrAlaGluGlySer-Leu-----TyrAsnVa 40
DB 4 GTGATGGAATATGCAGAGGGGGCTCATGTAATGTTGTCGCTTCTTCAAGT 63
OY 40 lleuhsiglyalagluProleuprotyrThrAlaahisAlaMetSertpCysLe 60
DB 64 GCTGATGCTGCTGAACCATTCCTTACTACACTGCTCTCATGCATGAGCTGTTT 123
OY 60 uclnlysserlnglyValAlaTyrlleuhsisermetGlnProlysalaleuilehisAr 80
DB 124 ACAGTGTTCACAGAGAGTGTCTTACCTGCACAGCATGAGCCCAAGGCTGATTCAAG 183
OY 80 GASpleuysProProAsnleuLeuValAlaGlyGlyThrValleuLysileCysAs 100
DB 164 GGACCTCAAGCTCCAACTTCTGCTGTCAGAGAGGACAGTCTTAAAAATCTGGA 243
OY 100 pPheGlyThrAlaCysAspIleGlnThrHisermThrAsnlysglySerrAlaIaTr 120
DB 244 TTTTGTACAGCTGTGACATCAACACACATGACCAATATAAGGAGTGTGCTTG 303
OY 120 pMetAlaProGluValPheGluGlySerAsnTyrsertGlyLysCysAspValPheSertR 140
DB 304 GATGGCGCTGAAGTGTGAAGGTAGCAATTAACAGTAAGAGTGTGCTTCAGCTG 363
OY 140 pGlyIlelleuLeuTyrGluValIle-ThrArgArgLysProPheAspGluIleGlyLP 160
DB 364 GGGTATTATCCCTCTGGAGATGATATAACCGCGGAAACCTTCATGATCGTGGCC 423
OY 160 roAlaPheArgIleMetTyrAlaValHisAsnGlyThrArgProProleuileLysAsnL 180
DB 424 CAGCTTTCAGATCATGTGGGCTTTCATTAATGACATGACCAACCACTGATCAAAAATT 483
OY 180 euProlyProIleGluSerLeuMetThrArgCysTyrSerLysAspProserGlnArgP 200
DB 484 TACCTAAGCCATTGAGAGCTTGATGACCGCTG-TGGTCTAAGAGACCCATCTCAGCG-C 541
OY 200 roSerMetGluGluIleValIysileMetThrHisLeuMetArgTyrPheProGlyAlaA 220
DB 542 CTTCAATGAGAGAAATTGTGAAAAATATGACTCACTTATCGGATCTTCCAGAGACGG 601
OY 220 spGluProleuGlnTyrProCysGln 228
DB 602 ATGACGA-TTACAGTATACTTGTCTGAG 626
```

RESULT 8

AL525728

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Genoscope - Centre National de Sequencage

BP 191 91006 Evry cedex - France

Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr.

Location/Qualifiers

1..998

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_image="CS0DC013YE20"

/clone_lib="LTI_NFL003_NBC3"

/sex="male"

/tissue_type="neuroblastoma cells"

/lab_host="DH10B"

/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA

was primed with a NotI-Oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pCMVSPORT 6

vector. Library was normalized. Library was constructed

by Life Technologies. Contact: Feng Liang Life

Technologies, a division of Invitrogen 9800 Medical Center

Drive Rockville, Maryland 20850, USA Fax: (1) 301 610

8371 Email: fliang@lifetech.com URL:

http://fulllength.invitrogen.com"

BASE COUNT 245 a 225 c 268 g 247 t 13 others

ORIGIN

Alignment Scores:

Pred. No.:	3.58e-97	Length:	998
Score:	932.50	Matches:	188
Percent Similarity:	95.43%	Conservative:	0
Best Local Similarity:	95.43%	Mismatches:	6
Query Match:	74.48%	Indels:	6
DB:	9	Gaps:	0

US-09-830-144-2_COPY_76_303 (1-228) x AL525728 (1-998)

```
OY 1 ValGluLeuArgGln-LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAl 20
DB 414 GTAGACTTCGGAGAGATTATCCGCTGTGAACATCTTAATATTGTAAGCTTATGAGAC 473
OY 20 aCysLeuAsnProValCysLeuValMetGluTyr-AlaGluGlyGlySerLeuTyraAnV 40
DB 474 CTGCTTGAATCCAGTGTGCTTGTGATGGAATATAGCTGAAGGGGCTCTTATATAATG 533
OY 40 alleuhsiglyalagluProleuprotyrThrAlaahisAlaMetSertpCysL 60
DB 534 TGCTGATGCTGCTGAACCATTCATTAATTAATGCTGCCACCAATGATGTGTT 593
OY 60 euGlnCysSerGlnGlyValAlaTyrlleuhsisermetGlnProlysalaleuilehisA 80
DB 594 TTAGAGTTCACCAAGAGAGGCTTATCTTCACAGATGCAACCAAGGCTTAATTCACA 653
OY 80 rGAspleuysProProAsnleuLeuValAlaGlyGlyThrValleuLysileCysA 100
DB 654 GGGACCTGGAACACCAACTTACTGCTGTGCAAGGGGGGACAGTCTTAAAAATTTGTG 713
OY 100 spPheGlyThrAlaCysAspIleGlnThrHisermThrAsnlysglySerrAlaIaIaTr 120
```



```

Db 714 A T T T T G T A C A G C T G T G A C A T T C A C A C A C A T G A C C A A T A A C A A G G G A G T G T G C T T 773
Qy 120 r p M e t A l a P r o G l u V a l P h e G l u G l y S e r A s n T y r S e r G l u L y s C y s A s p V a l P h e S e r T 140
Db 774 G G A T G C C A C C T G A A G T T T T T G A A G T A G T A A C A G T G A A A A A T G T G A C G T T T C A G C T 833
Qy 140 r p G l y l e i l e L e u T r p G l u V a l l e T h r A r g A r g P r o P h e A s p G l u l e G l y G l y P 160
Db 834 G G G G T A T T A T T C T T T G G G A A G T G A T A M G - C G T C G G A A C C T T T - C A T G A G A T T G T G G C C 891
Qy 160 r o A l a P h e A r g l e M e t T r p A l a V a l - H i s A s n G l y T h r A r g P r o P r o L e u l e l y s A s n 179
Db 892 C A G C T T T C C A A T C A T G T G G C T G T T C A T A T G T T A C T C G A C C A C C A C T G A T T A A A A A T 951
Qy 180 L e u P r o L y s P r o l e u S e r l e u M e t T h r A r g C y s T r p S e r L y s 194
Db 952 T T A C C T A A G C C A T T - G A G A G C C T G A T G A C T C G T T G T G T Y T A A A 995

RESULT 9
AW960377
LOCUS EST372448 MAGe resequences, MAGF Homo sapiens cDNA, mRNA sequence. EST 01-JUN-2000
DEFINITION AW960377
ACCESSION AW960377.1 GI:8150061
VERSION EST.
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 542)
AUTHORS Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaspar, R., Gay, C., Holt
I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and
Quackenbush, J.
TITLE Assessment of gene expression patterns in a model of colon tumor
metastasis using a 19,200 element cDNA microarray
JOURNAL Unpublished (2000)
COMMENT Contact: John Quackenbush
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: johnq@tigr.org
Plate: 145
Seq primer: Reverse.

FEATURES
source
1..542
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="MAGE resequences, MAGF"
/note="Vector: pBluescriptSKm"
BASE COUNT 149 a 102 c 136 g 155 t
ORIGIN

Alignment Scores:
Pred. No.: 1,08e-87 Length: 542
Score: 847.00 Matches: 159
Percent Similarity: 97.55% Conservatives: 0
Best Local Similarity: 97.55% Mismatches: 3
Query Match: 67.65% Indels: 1
DB: 10 Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x AW960377 (1-542)

Qy 1 V a l G l u L e u A r g G l n L e u S e r A r g V a l A s n H i s P r o A s n l l e V a l l y s L e u T y r G l y A l a 20
Db 51 G T A G A G C T T C C G A G T A T C C C G T G T G A C C A C C A T C C T A A T A T T G T A A G C T T T A T G A G C C 110
Qy 21 C y s L e u A s n P r o V a l C y s L e u V a l M e t G l u T y r A l a C l u G l y S e r L e u T y r A s n V a l 40
Db 111 T G C T T G A A T C C A G T G T C T T G T G A T G A A T A T G C T G A A G G G G C T C T T T A T A T A T G T 170
Qy 41 L e u H i s G l y A l a G l u P r o L e u P r o T y r T y r T h r A l a H i s A l a M e t S e r T r p C y s L e u 60

```

```

Db 171 C T G C A T G T G T C T G A A C C A T T G C C A T A T T A C T G C T G C C C A C G C A A T G A T T G T G T T T A 230
Qy 61 G l n C y s S e r G l n G l y V a l l a l a T y r L e u H i s S e r M e t G l n P r o L y s A l a L e u l l e H i s A r g 80
Db 231 C A G T G T T C C C A A G A G T G G C T T A T C T T C A G A C A T G C A A C C C A A A G C G C T A A T T C A C A G G 290
Qy 81 A s p L e u l y s P r o A s n L e u L e u V a l A l a G l y G l y T h r V a l L e u l y s l l e C y s A s p 100
Db 291 G A C C T G A A C C A C C A A A C T T A C T G C T G T G C A G G G G G A C A G T T C T A A A A T T T G T G A T 350
Qy 101 P h e G l y T h r A l a C y s A s p l l e G l n T h r H i s M e t T h r A s n A s n L y s G l y S e r A l a A l a T r p 120
Db 351 T T T G G T A C A G C C T G T G A C A T T C A G A C A C A C A T G A C C A A T A A C A A G G G G A G T G C T G T T G G 410
Qy 121 M e t A l a P r o G l u V a l P h e G l u G l y S e r A s n T y r S e r G l u L y s C y s A s p V a l P h e S e r T r p 140
Db 411 A T G G C A C C T G A A G T T T T T G A A G G T A G T A A T T A C A G T G A A A A A I G T G A C G T C T T C A C G T G G 470
Qy 141 G l y l e i l e L e u T r p G l u V a l l e T h r A r g A r g - L y s P r o P h e A s p G l u l e G l y G l y P r 160
Db 471 G G T A T T A T T C T T T T G G A A G T G A T A A C G C G T C G G A A C C C T G T G A T G A G A T T G T G G C C C 530
Qy 160 o A l a P h e 162
Db 531 A A C T T T C 537

RESULT 10
BQ386875 641 bp mRNA linear EST 22-MAY-2002
LOCUS NISC mn20f09.y1 NICHHD XGC Ov1 Xenopus laevis cDNA clone
DEFINITION IMAGE:5073953 5', mRNA sequence.
ACCESSION BQ386875
VERSION BQ386875.1 GI:21074562
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus.
1 (bases 1 to 641)
NIH-XGC http://image.llnl.gov/image/hml/xenopuslib.info.shtml.
National Institute of Child Health and Human Development, National
Cancer Institute, Xenopus Gene Collection
Unpublished (2002)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
cDNA Library Preparation:
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC)
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
info@image.llnl.gov
Plate: LLAM1198 row: L column: 18
Seq primer: M13RP1 reverse primer (ABI).
FEATURES
source
1..641
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone_lib="IMAGE:5073953"
/clone_lib="NICHHD XGC Ov1"
/sex="female"
/lab_host="DH10B (phage-resistant)"
/note="Organ: ovary; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.0 kb. Constructed by Life
Technologies."
BASE COUNT 200 a 134 c 139 g 168 t
ORIGIN

Alignment Scores:
Pred. No.: 8.45e-85 Length: 641

```

Score: 823.00 Matches: 150
 Percent Similarity: 98.05% Conservative: 1
 Best Local Similarity: 97.40% Mismatches: 3
 Query Match: 65.73% Indels: 0
 DB: 14 Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x BQ386875 (1-641)

QY 75 LysAlaLeuIleHisArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThr 94
 Db 13 AAGCTCTATTACAGAGACCTCAACCACTTCCTGCTGAGTGGAGGACC 72
 QY 95 ValLeuLysIleCysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsn 114
 Db 73 GTTCTTAAGATTGTGACTTGGTACAGCCTGGATATTCAAGCTCACATGACTATAAC 132
 QY 115 LysGlySerAlaAlaTrpMetAlaProGluValPheGlyGlySerAsnTyrSerGluLys 134
 Db 133 AAGGAAAGTCAGACATGATGGCTCCAGAAATTTTGAAGGTAGCAATTACAGCAGAAA 192
 QY 135 CysAspValPheSerTrpGlyIleIleLeuTrpGluValIleThrArgArgLysPhe 154
 Db 193 TGTGAGATTATTAGTTGGGAAATATTCTTTGGAGATATAACCGAAGAAACCTTTC 252
 QY 155 AspGluIleGlyGlyProAlaPheArgIleMetTrpAlaValHisAsnGlyThrArgPro 174
 Db 253 GATGAATAATGGTGTCCACAGTTCCGTATATGTGGGCTTTCACATGGTACTGCCCA 312
 QY 175 ProLeuIleLysAsnLeuProLysProIleGlySerLeuMetThrArgCysTrpSerLys 194
 Db 313 CCATTAAATTAATAAATTTGCTTAAGCTTATGAAGCTTAATGACTCGCTGCTGCAAA 372
 QY 195 AspProSerGlyArgProSerMetGluGluIleValLysIleMetThrHisLeuMetArg 214
 Db 373 GATCCGCCCAAGACCTTCAATGAGAGATTGTCACATATGATCATCTAATGACAG 432
 QY 215 TyrPheProGlyAlaAspGluProLeuGlnTyrProCysGln 228
 Db 433 TATTTTCTGGAGCGGATGTTCTTACAGTATCTTGTTCAG 474

RESULT 11

AL550589 827 bp mRNA linear EST 16-FEB-2001
 LOCUS AL550589 LTI NFL006.PL2 Homo sapiens cDNA clone CS0D1058Y014 5
 DEFINITION prime, mRNA sequence.

ACCESSION AL550589
 VERSION AL550589.1 GI:12887700
 KEYWORDS EST.

SOURCE

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 827)
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

source

1..827
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0D1058Y014"
 /clone_id="LTI NFL006.PL2"
 /tissue_type="Placenta"
 /note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
 was primed with a NotI-oligo(dT) primer. Five prime end
 enriched, double-stranded cDNA was digested with Not I and
 cloned into the Not I and Eco RV sites of the pCMVSPORT 6
 vector. Library was normalized. Library was constructed by
 Life Technologies. Contact : Feng Liang Life Technologies,

a division of Invitrogen 9800 Medical Center Drive
 Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
 Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/

BASE COUNT 196 a 178 c 228 g 209 t 16 others

ORIGIN

Alignment Scores:

Pred. No.: 1,85e-80 Length: 827
 Score: 787.00 Matches: 147
 Percent Similarity: 98.66% Conservative: 0
 Best Local Similarity: 98.66% Mismatches: 1
 Query Match: 62.86% Indels: 1
 DB: 9 Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x AL550589 (1-827)

QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20
 Db 382 GTAGAGCTTGGCAGATTATCCCTGTRAACCATCTTAATTTGTAAGCTTTATGAGCC 441
 QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGlyGlySerLeuTyrAsnVal 40
 Db 442 TGCCTGAATCAGATGTGTCTTGTGATGAAATATGCTGAAGGGGCTCTTATATATGTR 501
 QY 41 LeuHisGlyAlaGluProLeuProTyrTyrTrpAlaAlaHisAlaMetSerTrpCysLeu 60
 Db 502 CTGCATGTGTCTGAACCATTTGCCATATTAATCTGCTGCCACGCAATGATGGTGTTTA 561
 QY 61 GlnCysSerGln-GlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80
 Db 562 CAGGTTCACCAAGAGTGGCTTATCTTCACAGACATGCCAACCAAGCCCTAATTCACAG 621
 QY 80 GAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsn 100
 Db 622 GGAACCTGAAACCAACAACTTACTGCTGTGTCAGGGGAGACATCTTAATAATTGTGA 681
 QY 100 pPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaIle 120
 Db 682 TTTTGTACAGCTGTGACCTTGAACACATGACCAATTAACAAGGGAGTGTGCTTGG 741
 QY 120 pMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTr 140
 Db 742 GATGGCACCTCGAAGTTTGTGAAGTATTAATTAACGTGAAAAATGTGACCTTTCACGCT 801
 QY 140 pGlyIleIleLeuTyrGluValIle 148
 Db 802 GGGTATTATTCTTTGGGAAGTGATA 826

RESULT 12

BQ219348 1006 bp mRNA linear EST 02-MAY-2002
 LOCUS BQ219348 1006 bp mRNA linear EST 02-MAY-2002
 DEFINITION AGENCOURT 7260864 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5785294
 5', mRNA sequence.

ACCESSION BQ219348
 VERSION BQ219348.1 GI:20400748
 KEYWORDS EST.

SOURCE

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1006)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs@mail.nih.gov

Tissue Procurement: ATCC
 cDNA library Preparation: Life Technologies, Inc.
 cDNA library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:

http://image.llnl.gov
Plate: LLAM12872 row: c column: 23
High quality sequence stop: 624.
FEATURES
source
1. .1006
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5785294"
/clone_lib="NIH MGC 71"
/tissue type="leiomyosarcoma"
/lab host="DH10B (phage-resistant)"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 2.1 kb. "
BASE COUNT 240 a 240 c 293 g 233 t
ORIGIN

Alignment Scores:
Pred. No.: 2,16e-76 Length: 1006
Score: 753.00 Matches: 151
Percent Similarity: 78.39% Conservative: 5
Best Local Similarity: 75.88% Mismatches: 17
Query Match: 60.14% Indels: 27
DB: 14 Gaps: 4

US-09-830-144-2_COPY_76_303 (1-228) x BQ219348 (1-1006)

QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20
|||
Db 388 GTAGAGCTTCGGAGTTATCCGGTGAACATCCTTAATTTGAAGCTTTATGAGCC 447
|||
QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnVal 40
|||
Db 448 TGCTTGAATCCAGTGTCTCTGTGATGAATATGCTGAAGGGGCTCTTTATATATG 507
|||
QY 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 60
|||
Db 508 CTGCATGTGCTGAACCATTTGCCATATTATCTGCTGCCACGCAATAGTTGCTTTTA 567
|||
QY 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80
|||
Db 568 CAGTGTTCACAGAGTGGCTTATCTTCAGCATGCAACCAAGCGCTTAATTCAGG 627
|||
QY 81 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100
|||
Db 628 GACCTGAACACACCACTTACTGCTGTTGCAGGGGGACAGTTCT-AAAATTTGTGAT 686
|||
QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 120
|||
Db 687 TTGGGTACAGCCTGTGACTTTCAGACACATGACCAATTTACAAGGGGAGTGTGCTTGG 746
|||
QY 121 MetAlaProGluValPheGluGlySerAsnTyrSerGlnLysCysAspValPheSerTrp 140
|||
Db 747 ATGCACCTGGAAATTTTGGAGGAGTAACCCGTCGGGAAACCTTTTGNAGAGAAATTGGGG 806
|||
QY 141 GlyIleIleLeuTrpGluValIleThrArgArg---LysProPheAsp---GluIleGly 158
|||
Db 807 GGGAAATTTCTTTGGGAAAGATAACCCGTCGGGAAACCTTTTGNAGAGAAATTGGGG 866
|||
QY 159 GlyProAlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLys 178
|||
Db 867 GGGCAACTTTT-----878
|||
QY 179 AsnLeuProLysProIle-----GluSerLeuMetThrArgCys 191
|||
Db 879 -----CCAAACCACTGGGGCGGGGCTCATACATGATGACCCCAACCCCGCTGC 929
|||
RESULT 13
BG548917 602575750F1 NIH MGC_77 Homo sapiens cdna clone IMAGE:4703580 5',
LOCUS mRNA sequence.
DEFINITION BG548917
ACCESSION

BG548917.1 GI:13547582
EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 686)
NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: rcapbs-@email.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cdna Library Preparation: CLONTECH Laboratories, Inc.
cdna Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCMI542 row: d column: 13
High quality sequence stop: 671.
FEATURES
Location/Qualifiers
source
1. .686
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4703580"
/clone_lib="NIH MGC 77"
/lab host="DH10B (TI phage-resistant)"
/note="Organ: lung; Vector: pBNR-LIB (Clontech); Site 1:
SfiI (ggcgccctcgcc); Site 2: SfiI (ggccattggcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CAGGCCATTATGGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCCGAGCGCGGCATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.9
kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH MGC Library."

BASE COUNT 202 a 143 c 169 g 171 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 4.5e-76 Length: 686
Score: 748.00 Matches: 144
Percent Similarity: 95.39% Conservative: 1
Best Local Similarity: 94.74% Mismatches: 4
Query Match: 59.74% Indels: 3
DB: 12 Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x BG548917 (1-686)

QY 80 ArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCys 99
|||
Db 4 AGGGACCTTGAACACCACTTACTGCTGTAGTGGGGGACAGTTCTAAAAATTTGT 63
|||
QY 100 AspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAla 119
|||
Db 64 GATTTTGTGTACAGCTGTGACATTCAGACACATGACCAATAACAGGGAGTGTGCT 123
|||
QY 120 TrpMetAlaProGluValPheGluGlySerAsnTyrSerGlu-LysCysAspValPheSe 139
|||
Db 124 TGGATGGCACCTGAAGTTTGTGAAGGTAGTAATTTACAGTGAACAATGTGACGCTTCAG 183
|||
QY 139 rTrpGlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGl 159
|||
Db 184 CTGGGGTATTATCTTTTGGGAAGTGATAACCGCTCGGAAACCTTTGATGAGATTGGTGG 243
|||
QY 159 yProAlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLys-A 179
|||
Db 244 CCCAGCTTTCGAATCATGTGGGCTGTTCAATGTTACTCGNACAGCACTGATAACAA 303
|||
QY 179 snLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGln- 198
|||

```

Db 304 ATTACTTAAGCCCATTTGAGAGCCTGATGACTGTGTTGTCCTTAAGATCCTCCAG 363
Qy 199 AagpSerMetGluGluIleValIysIleMetThrHisLeuMetArgTyrPheProGly 218
Db 364 CGCCCTTCATGAGGAAATTTGAAAATTAATGACTCACTGTAGCGGTACTTCCAGGA 423
Qy 219 AlaAspGluProLeuGlnTyrProCysGln 228
Db 424 GCAGATGAGCCATTACAGTATCCTGTGAG 453

RESULT 14
BI093821 771 bp mRNA linear EST 20-JUN-2001
LOCUS 602860601F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:5001504 5',
DEFINITION mRNA sequence.
ACCESSION BI093821
VERSION BI093821.1 GI:14512151
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE NIH-MGC http://mgs.nci.nih.gov/.
AUTHORS 1 (bases 1 to 771)
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM1035 row: b column: 01
High quality sequence stop: 768.
FEATURES
Location/Qualifiers
source 1..771
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5001504"
/clone_1lb="NIH_MGC_10"
/cell_line="MGC36"
/lab_host="DH10B"
/notes="Organ: cervix; Vector: pCMV-SPORT6; Site: 1; NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5 kb. Library prepared by Life
Technologies."
BASE COUNT 230 a 163 c 187 g 191 t
ORIGIN
Alignment Scores:
Pred. No.: 3,49e-75 Length: 771
Score: 741.00 Matches: 133
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 59.19% Indels: 0
DB: 13 Gaps: 0
US-09-830-144-2_COPY_76_303 (1-228) x BI093821 (1-771)
Qy 96 LeuLysIleCysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnLys 115
Db 1 CTTAAATTTTGATTTTGGTACAGCCTGTGACATTCACACACATGACCAATACAG 60
Qy 116 GlySerAlaAlaIlePheMetAlaProGluValIlePheGluGlySerAsnTyrSerGluLysCys 135
Db 61 GGGAGTGTGCTTGATGACCTGACCTGAAGTTTGAAGTAGTAATTAACAGTAAATAATGT 120
Qy 136 AspValPheSerTyrGlyIleIleLeuTyrGluValIleThrArgArgLysProPheAsp 155
Db 121 GACGCTTCAGTGGGGGTATATCTTTGGGAAAGTATACCGGTGGAAACCTTTGAT 180

```

```

Qy 156 GluIleGlyIleProAlaPheArgIleMetTyrPalaValHisAsnGlyThrArgProPro 175
Db 181 GAGATGTGTGGCCCGACGCTTCCGATCATGTGGCTGTTTCAATATGACTCCACCACCA 240
Qy 176 LeuIleLysAsnLeuProLysProIleGluSerLeuMetThrArgCysTyrPheSerLysAsp 195
Db 241 CTGATATAAAATTTTACTTAAGCCCATTTGAGAGCCTGATGACTGTGTTGTTGTTAAAGAT 300
Qy 196 ProSerGlnArgProSerMetGluGluIleValIysIleMetThrHisLeuMetArgTyr 215
Db 301 CTTCCCGAGCGCCCTTCATGAGGAAATTTGAAAATTAATGACTCACTGTAGCGCGTAC 360
Qy 216 PheProGlyAlaAspGluProLeuGlnTyrProCysGln 228
Db 361 TTTCAGAGAGCATGAGCCATTACAGTATCCTGTGAG 399

RESULT 15
AV398933 682 bp mRNA linear EST 05-FEB-2000
LOCUS AV398933
DEFINITION Bombyx mori ovary BmNPV infected; 6 hr after inoculation
ACCESSION AV398933
VERSION AV398933.1 GI:6902585
KEYWORDS EST.
SOURCE domestic silkworm.
ORGANISM Bombyx mori
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Bombycoidea; Bombycidae; Bombyx.
REFERENCE 1 (bases 1 to 682)
AUTHORS Mita,K., Morimoto,M., Shimada,T., Okano,K. and Maeda,S.
TITLE Bombyx mori cDNA
JOURNAL Unpublished (2000)
COMMENT Contact: Mita K
Genome Research Group
National Institute of Radiological Sciences
Anagawa 4-9-1, Inage, Chiba 263-8555, Japan
Email: kmkita@nirs.go.jp
method:uni-directional, sequence direction:sequenced from T3 primer
(5' -> 3')
Project="Silkworm Genome Program in MAF, and Research for the
Future Program in JSPS". see 'SilkBase',
<http://www.ab.a.u-tokyo.ac.jp/silkbase/>, for whole ESTdb.
FEATURES
Location/Qualifiers
source 1..682
/organism="Bombyx mori"
/db_xref="taxon:7091"
/clone="NV060441"
/clone_1lb="Bombyx mori ovary BmNPV infected; 6 hr after
inoculation"
/tissue_type="ovary"
/cell_type="BmNPV infected cell"
/dev_stage="BmNPV infected; 6 hr after inoculation"
BASE COUNT 200 a 130 c 191 g 160 t
ORIGIN
Alignment Scores:
Pred. No.: 8.37e-75 Length: 682
Score: 737.00 Matches: 139
Percent Similarity: 85.80% Conservative: 12
Best Local Similarity: 78.98% Mismatches: 21
Query Match: 58.87% Indels: 4
DB: 10 Gaps: 3
US-09-830-144-2_COPY_76_303 (1-228) x AV398933 (1-682)
Qy 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20
Db 159 ATCAGAGTACGCCGATATCTCGGTATGTGCACCAAACTTGTGGATTTGTCGAGGCG 218
Qy 21 CysLeuAsn-----ProValCysLeuValMetGluTyrAlaGluGlySerLeuTyr 38

```

```
Db 219 TGCACGCGGGGTCTCACGTATGCCTCGTCATGGAGTATGCAGAAGGTGGCTCTCTCTAT 278
QY 39 AsnValLeuHisGlyAlaGluProLeuProTyrThrAlaAlaHisAlaMetSerTrp 58
Db 279 AATGTACTACACAAC---CGACCCAAACCCAAATACACAGCAGCACATGCAATGAGCTGG 335
QY 59 CysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIle 78
Db 336 GCCCGCCAGTGGCTGAGGGTGTAGCATATCTTCACTCCATGAACCAAAACCATTAATA 395
QY 79 HisArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIle 98
Db 396 CACAGGGATTTAAACCAACCACTATTATTATTAGTGGGTGGAGGCAACGGCTGAAGATT 455
QY 99 CysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAla 118
Db 456 TGTGATTTTGGCACAGCAGCTGATAGGCAACATACATGACCAATATAAAGGCAGTGCT 515
QY 119 AlaTrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPhe 138
Db 516 GCTTGGATGGCTCTGAGGTATTGAAGGATCGACATACACAGAAAAATGTGATGTGTC 575
QY 139 SerTrpGlyIleIleLeuTrpGluValIleThrArgLysProPheAspGluIleGly 158
Db 576 TCATGGGGGATAATACTGTGGGAGGTATTATCTAGGAGGAAACCATTTGAAGAA---GGT 632
QY 159 GlyProAlaPheArgIleMetTrpAlaValHisAsnGlyThrArgPro 174
Db 633 GGTTCGCATATAGATAATGTGGCGGTGCATACAGGTCNAAGACCA 680
```

Search completed: December 10, 2002, 05:34:37
Job time : 2223 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 10, 2002, 03:59:35 ; Search time 68 Seconds

(without alignments)
1028.269 Million cell updates/sec

Title: US-09-830-144-2_COPY_76_303

Perfect score: 1252

Sequence: 1 VELRQLSRVNHPTIKLYGA.....MTHLMRYFGADELQVPCQ 228

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlp
-Q/cgn2_1/USPTO.spool/US09830144/runat_04122002_141755_4367/app_query.fasta_1.391
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORW=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09830144@cgn_1_1_31@runat_04122002_141755_4367 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:
1: /cgn2_6/prodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/prodata/1/ina/PCTUS.COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1252	100.0	1788	4	US-09-529-279-14 Sequence 14, Appli
2	1252	100.0	2856	2	US-08-685-625A-5 Sequence 5, Appli
3	1252	100.0	2856	4	US-09-529-279-3 Sequence 3, Appli
4	1247	99.6	2443	3	US-08-685-625A-1 Sequence 1, Appli
5	371	29.6	1365	3	US-09-221-235-6 Sequence 6, Appli
6	371	29.6	1365	3	US-09-221-928-6 Sequence 6, Appli
7	371	29.6	1365	3	US-09-221-527-6 Sequence 6, Appli
8	371	29.6	1365	3	US-09-221-236-6 Sequence 6, Appli
9	371	29.6	1365	3	US-09-221-416-6 Sequence 6, Appli
10	371	29.6	1365	4	US-09-221-245-6 Sequence 6, Appli
11	371	29.6	1365	4	US-09-163-115-6 Sequence 6, Appli
12	371	29.6	1365	4	US-09-221-528-6 Sequence 6, Appli

13	371	29.6	1365	4	US-09-593-553-6	Sequence 6, Appli
14	371	29.6	1365	4	US-09-221-237-6	Sequence 6, Appli
15	371	29.6	2120	3	US-09-221-235-4	Sequence 4, Appli
16	371	29.6	2120	3	US-09-221-928-4	Sequence 4, Appli
17	371	29.6	2120	3	US-09-221-527-4	Sequence 4, Appli
18	371	29.6	2120	3	US-09-221-236-4	Sequence 4, Appli
19	371	29.6	2120	3	US-09-221-416-4	Sequence 4, Appli
20	371	29.6	2120	4	US-09-221-245-4	Sequence 4, Appli
21	371	29.6	2120	4	US-09-163-115-4	Sequence 4, Appli
22	371	29.6	2120	4	US-09-221-528-4	Sequence 4, Appli
23	371	29.6	2120	4	US-09-593-553-4	Sequence 4, Appli
24	371	29.6	2120	4	US-09-221-237-4	Sequence 4, Appli
25	353.5	28.2	3389	1	US-08-395-580-1	Sequence 1, Appli
26	353.5	28.2	3426	1	US-08-205-018-1	Sequence 1, Appli
27	338.5	27.0	2505	4	US-09-291-839-3	Sequence 3, Appli
28	338.5	27.0	3025	4	US-09-291-839-1	Sequence 1, Appli
29	323	25.8	2890	1	US-07-928-464-1	Sequence 1, Appli
30	323	25.8	2890	5	PCT-US93-07347-1	Sequence 1, Appli
31	323	25.8	3033	1	US-08-003-311B-1	Sequence 1, Appli
32	323	25.8	3033	1	US-08-261-432-1	Sequence 1, Appli
33	305.5	24.4	3516	3	US-09-188-930-257	Sequence 257, Appl
34	304.5	24.3	4508	5	PCT-US93-06251-34	Sequence 34, Appl
35	294	23.5	1888	3	US-09-188-930-66	Sequence 66, Appl
36	293.5	23.4	2574	4	US-08-142-529-2	Sequence 2, Appli
37	280	22.4	1875	1	US-08-070-165F-3	Sequence 3, Appli
38	280	22.4	1875	2	US-08-885-418-3	Sequence 3, Appli
39	278	22.2	933	2	US-08-701-191A-4	Sequence 4, Appli
40	278	22.2	1056	2	US-08-701-191A-5	Sequence 5, Appli
41	278	22.2	2469	1	US-07-997-133-2	Sequence 2, Appli
42	278	22.2	2469	5	US-07-997-133-2	Sequence 2, Appli
43	278	22.2	2662	2	US-08-451-822A-14	Sequence 14, Appl
44	278	22.2	2662	4	US-08-323-430-14	Sequence 14, Appl
45	275	22.0	2469	1	US-08-459-296-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-09-529-279-14

; Sequence 14, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278 US/09/529, 279
; CURRENT FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 1788
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)..(1776)
US-09-529-279-14

Alignment Scores:
Pred. No.: 2.3e-140 Length: 1788
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x US-09-529-279-14 (1-1788)

```

QY      1 ValGluLeuAArgGluLeuSerArgValaAsnHisProAsnIleValLysLeuTyrgIyala 20
Db      232 GTAGAGCTTCGGGAGATTATCCCGTGAACCACTTAATATGTAAGCTTATATGAGGCC 291
QY      21 CysLeuAsnProValCysLeuValMetGluTyraIaGlyGlySerLeuTyraAsnVal 40
Db      292 TGGTTGAATCCAGTGTGTCTTGTGATGGAATATGCTGAAGGGGGCTCTTATATATATGTG 351
QY      41 LeuHisGlyAlaGluProLeuProTyryTyThraAlaHisAlaMetSerTyrcysLeu 60
Db      352 CTCGATGAGTCTGAACCATGTCATATTAATGCTGCCACGCAATGAGTGTGCTTTA 411
QY      61 GlncysSerGlnGlyValaIaTyrlLeuHisSerMetGlnProLysAlaLeuIleHisarg 80
Db      412 CAGTGTTCCTCCAAAGAGTGGCTTATCTTCACAGATGCAACCCAAACGGCTTAATTCACAG 471
QY      81 AspLeuLysProProAsnLeuLeuValaIaGlyGlyThraValLeuLysIleCysAsp 100
Db      472 GACCTGAACCAACCAACTTACTGCTGCTGGTTCAGGGGGGACAGTCTTAATAATTTGTGAT 531
QY      101 PheGlyThraAlaCysAspIleGlnThraHisMetThraAsnLysGlySerAlaIaIaIaIa 120
Db      532 TTGGTACAGCCTGTGACATTCAGACACATGACCAATTAACAGGGGAGTGTCTGTGG 591
QY      121 MetAlaProGluValaPheGluGlySerAsnTyrySerGluLysCysAspValaPheSerTrp 140
Db      592 ATGGCAGCTGAAGATTTTGAAGTAGTAATTAACAGTGAATAATGTCACCTTTCAGCTGG 651
QY      141 GlyIleIleLeuTyrgIyValaIleThraArgLysProPheAspGluIleGlyGlyPro 160
Db      652 GGTATATTTCTTTGGGAAGTGAATACGCTCGAAACCTTTGATGAGATTTGTGGCCCA 711
QY      161 AlaPheArgIleMetTrpAlaValHisAsnGlyThraArgProProLeuIleLysAsnLeu 180
Db      712 GCTTTCGATCATGTGGGCTGTTCATATGTAATGTAATCGACCACTGATTAATAATTTA 771
QY      181 ProLysProIleGluSerLeuMetThraArgCysTrpSerLysAspProSerGlnArgPro 200
Db      772 CCAAGCCCATGAGAGCCTGATGAGCTCGTGTGTCTTAAGATCTTCCACGAGCCCT 831
QY      201 SerMetGluGluIleValLysIleMetThraHisLeuMetArgTyryPheProGlyAlaAsp 220
Db      832 TCATGAGAGAAATGTGAAATAATGACTCACTTGATCGGTACTTTCAGAGACAAAT 891
QY      221 GluProLeuGlnTyryProCysGln 228
Db      892 GAGCCATTACAGTATCTGTGAG 915

RESULT 2
; Sequence 5, Application US/08685625A
; Patent No. 5945301
; GENERAL INFORMATION:
; APPLICANT: UNENO, Naoto
; APPLICANT: MATSUMOTO, Kunihiko
; APPLICANT: IRIE, Kenji
; TITLE OF INVENTION: NOVEL KINASE IN TGF- $\beta$  FAMILY SIGNAL
; TRANSDUCTION SYSTEM
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/08/685,625A
; FILING DATE: 24-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-253549
; FILING DATE: 29-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Meuth, Donna M.
; REGISTRATION NUMBER: 36,607
; REFERENCE/DOCKET NUMBER: 001560-267
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2656 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 183..1922
; US-08-685-625A-5

Alignment Scores:
Pred. No.: 4,17e-140 Length: 2656
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x US-08-685-625A-5 (1-2656)

QY      1 ValGluLeuAArgGluLeuSerArgValaAsnHisProAsnIleValLysLeuTyrgIyala 20
Db      408 GTAGAGCTTCGGGAGATTATCCCGTGAACCACTTAATATGTAAGCTTATATGAGGCC 467
QY      21 CysLeuAsnProValCysLeuValMetGluTyraIaGlyGlySerLeuTyraAsnVal 40
Db      468 TGGTTGAATCCAGTGTGTCTTGTGATGGAATATGCTGAAGGGGGCTCTTATATATATGTG 527
QY      41 LeuHisGlyAlaGluProLeuProTyryTyThraAlaHisAlaMetSerTyrcysLeu 60
Db      528 CTGATGTGTCTGAACCATTCATTAATTAATGCTGCCACGCAATGAGTGTGTTA 587
QY      61 GlncysSerGlnGlyValaIaTyrlLeuHisSerMetGlnProLysAlaLeuIleHisarg 80
Db      588 CAGTGTTCCTCCAAAGAGTGGCTTATCTTCACAGCATGCAACCCAAACGGCTTAATTCACAG 647
QY      81 AspLeuLysProProAsnLeuLeuValaIaGlyGlyThraValLeuLysIleCysAsp 100
Db      648 GACCTGAACCAACCAACTTACTGCTGCTGGTTCAGGGGGGACAGTCTTAATAATTTGTGAT 707
QY      101 PheGlyThraAlaCysAspIleGlnThraHisMetThraAsnLysGlySerAlaIaIaIaIa 120
Db      708 TTGGTACAGCCTGTGACATTCAGACACATGACCAATTAACAGGGGAGTGTCTGTGG 767
QY      121 MetAlaProGluValaPheGluGlySerAsnTyrySerGluLysCysAspValaPheSerTrp 140
Db      768 ATGGCAGCTGAAGATTTTGAAGTAGTAATTAACAGTGAATAATGTCACGCTTTCAGCTGG 827
QY      141 GlyIleIleLeuTyrgIyValaIleThraArgLysProPheAspGluIleGlyGlyPro 160
Db      828 GGTATATTTCTTTGGGAAGTGAATACGCTCGAAACCTTTGATGAGATTTGTGGCCCA 887
QY      161 AlaPheArgIleMetTrpAlaValHisAsnGlyThraArgProProLeuIleLysAsnLeu 180
Db      888 GCTTTCGATCATGTGGGCTGTTCATATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 947
QY      181 ProLysProIleGluSerLeuMetThraArgCysTrpSerLysAspProSerGlnArgPro 200

```


Db 948 CCTAAGCCATTGAGACCTGATGACTCGTTGGTCTAAAGATCCTCCCGCGCCT 1007
Qy 201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220
|||||
Db 1008 TCAATGGAGGAATTGTGAAATAATGACTCATTGATGGGTACTTTCCAGGAGCAGAT 1067
Qy 221 GluProLeuGlnTyrProCysGln 228
|||||
Db 1068 GAGCCATTACAGTATCCTTGTTCAG 1091
RESULT 3
US-09-529-279-3
; Sequence 3, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529,279
; CURRENT FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2656
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (183)..(1919)
US-09-529-279-3
Alignment Scores:
Pred. No.: 4.17e-140 Length: 2656
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 4
US-09-830-144-2_COPY_76_303 (1-228) x US-09-529-279-3 (1-2656)
Qy 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20
Db 408 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCATCCTAATATTGTAAAGCTTTATGGAGCC 467
Qy 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnVal 40
Db 468 TGGTTGAATCCAGTGTGCTTGTGATGGAAATATGCTGAAGGGGGCTCTTTATATATATGG 527
Qy 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 60
Db 528 CTGCATGGTGTCAACCATTTGCCATATTATCTGCTGCCACCAATGATGGTTTGA 587
Qy 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80
Db 588 CAGTGTTCACAGGAGTGGCTTATCTTCAGCATGCCAACCCAAAGCGCTAATTCACAGG 647
Qy 81 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100
Db 648 GACCTGAACACCACAACTTACTGTGTTGACGGGGGACAGTCTCTAAAAATTTGTGAT 707
Qy 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 120
Db 708 TTTGGTACAGCCGTGTGACATTACACACATGACACCAATCAACAGGGGAGTCTGCTTGG 767
Qy 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 140
|||||

Db 768 ATGGCACCTGAAGTATTTTGAAGTAGTAATTTACAGTGAATAATGTGACGTCTTCAGCTGG 827
Qy 141 GlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 160
|||||
Db 828 GGTATTATTCTTTGGGAAGTGATAACGCGTCGGAACCCCTTTGATGAGATTGGTGGCCCA 887
Qy 161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProLeuIleLysAsnLeu 180
|||||
Db 888 GCTTTCCGAATCATGTGGGCTGTTCAATAATGTTACTGACCACCACCTGATAAAAATTTA 947
Qy 181 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 200
|||||
Db 948 CCTAAGCCATTGAGAGCTGATGACTCGTTGGTCTAAAGATCCTTCCCGCGCCT 1007
Qy 201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220
Db 1008 TCAATGGAGGAATTGTGAAATAATGACTCATTGATGGGTACTTTCCAGGAGCAGAT 1067
Qy 221 GluProLeuGlnTyrProCysGln 228
|||||
Db 1068 GAGCCATTACAGTATCCTTGTTCAG 1091
RESULT 4
US-08-685-625A-1
; Sequence 1, Application US/08685625A
; Patent No. 5945301
; GENERAL INFORMATION:
; APPLICANT: UENO, Naoto
; APPLICANT: MATSUMOTO, Kunihiro
; APPLICANT: IRIE, Kenji
; TITLE OF INVENTION: NOVEL KINASE IN TGF-BETA FAMILY SIGNAL
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/685,625A
; FILING DATE: 24-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-253549
; FILING DATE: 29-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Meuth, Donna M.
; REGISTRATION NUMBER: 36,607
; REFERENCE/DOCKET NUMBER: 001560-267
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2443 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 157..1893
US-08-685-625A-1
Alignment Scores:
Pred. No.: 1.46e-139 Length: 2443

Score: 1247.00 Matches: 227
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.60% Indels: 0
DB: 2 Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x US-08-685-625A-1 (1-2443)

```
QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIysLeuTyrGlyAla 20
Db GTGGAGACTCCGGCAGTGTGTGCGCGTGAACCATCTTAACATTGTCAAGTTGTACAGAGCC 441
QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 40
Db TGCCTTAATCCAGTATGCTCTGTGATGGAATGACAGAGGGGGCTCATTTGTAATGTG 501
QY 41 LeuHisGlyValGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu 60
Db CTGCATGCTGCTGACCATTTGCTTACTACACTGCTGCTCATGSCCATGAGCTGGTGTTA 561
QY 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProIysAlaLeuIleHisArg 80
Db CACTGTTCCCAAGAGAGTGGCTTACTGCACAGCATGCACCCCAAGCGCTGATTACAGG 621
QY 81 AspleuIysProProAsnLeuLeuLeuValAlaGlyGlyThrValLeuIysIleCysAsp 100
Db GACCTCAAGCTTCCAACTTGCTGCTGTTGCGAGGAGGACAGTTCTTAATAATCTGCCAT 681
QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnIysGlySerAlaAlaTyr 120
Db TTTGGTACAGCTTGTGACATCCAAACACACATGACCAATATTAAGAGAGTGTGCTTGG 741
QY 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluIysCysAspValPheSerTyr 140
Db ATGGCGCTGCAAGTGTGTGAAGGTAGCAATTACAGTGAAGGTGATGCTTCACTGCTGG 801
QY 141 GlyIleIleLeuTyrGluValIleThrArgArgIysProPheAspIuIleGlyGlyPro 160
Db GGTATTATCTCTCGGAGAGATMACACGCCGGAACCCCTTCATAGATCGGTGACCCA 861
QY 161 AlaPheArgIleMetThrAlaValHisAsnGlyThrArgProProLeuIleIysAsnLeu 180
Db GCTTTCAAGATCATGTGGCTGCTTCATATGACACTGCACCACTGATCAAAATTTTA 921
QY 181 ProIysProIleGluSerLeuMetThrArgCysTyrPserIysAspProSerGlnArgPro 200
Db CTTAAGCCCATTTAGAGCTTGATGACACGCTGTGTCTAAGAACCATCTCAGCCGCT 981
QY 201 SerMetGluGluIleValIysIleMetThrHisLeuMetArgTyrPheProGluAlaAsp 220
Db TCAATGAGAGAAATTGTGAAATAATGACACCTGATGCGGTACTTCCAGAGGCGGAT 1041
QY 221 GluProLeuGlnTyrProCys 227
Db 1042 GAGCCATTACAGTATCTTGT 1062
```

RESULT 5
US-09-221-235-6
; Sequence 6, Application US/09221235
; Patent No. 6043040
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/221,235
; EARLIER FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1365
; TYPE: DNA

ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1..(1365)
US-09-221-235-6

Alignment Scores:

Pred. No.:	6,566-35	Length:	1365
Score:	371.00	Matches:	81
Percent Similarity:	56.28%	Conservative:	40
Best Local Similarity:	37.67%	Mismatches:	80
Query Match:	29.63%	Indels:	14
DB:	3	Gaps:	8

US-09-830-144-2_COPY_76_303 (1-228) x US-09-221-235-6 (1-1365)

```
QY 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIysLeuTyrGlyAlaCys 21
Db GTGGAGAGAAATCACTCACTGCTGCAGTACAGAAACATCATCCAGTTTATGAGATTAATT 216
QY 22 LeuAsnPro-----ValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsn 39
Db 217 CTTGAACCTCCCAACTATGACATTGTCCAGAAATATGCTTCTGTGGATCCTTATGAT 276
QY 40 ValLeuHisGly-----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSer 57
Db 277 TACATTAACTAACACAGAGTGAAGATG-----GATATGATCATCATTAATGACC 327
QY 58 TyrCysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProIysAlaLeu 77
Db 328 TGGGCACTGATGTAGCCAAAGGAATGATTAATTATGATGAGGCTCTGCAAGGTG 387
QY 78 IleHisArgAspleuIysProProAsnLeuLeuValAlaGlyGlyThrValLeuIys 97
Db 388 ATTCAACAGAGACCTCAAGTCAAGAAAGTTGTATTAACCTCTCATAGGA---GTACTGAAG 444
QY 98 IleCysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnIys----- 115
Db 445 ATCTGTACTTTGGT---GCCCTCGGTTCCATTAACATACACACACATGCTCTTGTT 501
QY 116 GlySerAlaAlaTyrMetAlaProGluValPheGluGlySerAsnTyrSerGluIysCys 135
Db 502 GGAACCTTCCCATGAGTGGCTCCAGAAAGTTATCCAGAGTCTCCCTGTGTCAGAAACTTGT 561
QY 136 AspValPheSerTyrGlyIleIleLeuTyrGluValIleThrArgArgIysProPheAsp 155
Db 562 GACACATATTCCTATGTGGTGTCTGTGGAGATGCTTAACAAGGAGGTCCCTTTTAA 621
QY 156 GluIleGlyGlyProAlaPheArgIleMetTyr---AlaValHisAsnGlyThrArgPro 174
Db 622 GGTITGAGAAGA-----TTACAAAGTACGTTGGCTTGTATGTGAAAAAAGAGATTA 675
QY 175 ProLeuIleIysAsnLeuProIysProIleGluSerLeuMetThrArgCysTyrPserIys 194
Db 676 ACCATTCCAGCAGTGGCCCGAGAAATTGCTGAATGTTATCATCATGCTGTGGAGAGCT 725
QY 195 AspProSerGlnArgProSerMetGluGluIleValIysIleMet 209
Db 736 GATGCCAAGAAACGGCCATCATTCACAAATCATTTCAATCTCG 780
```

RESULT 6
US-09-221-928-6
; Sequence 6, Application US/09221928
; Patent No. 6121030
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/221,928
; EARLIER FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 6

; LENGTH: 1365

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(1365)

US-09-221-928-6

Alignment Scores:

Pred. No.: 6.56e-35 Length: 1365
Score: 371.00 Matches: 81
Percent Similarity: 56.28% Conservative: 40
Best Local Similarity: 37.67% Mismatches: 80
Query Match: 29.63% Indels: 14
DB: 3 Gaps: 8

US-09-830-144-2_COPY_76_303 (1-228) x US-09-221-928-6 (1-1365)

QY 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCys 21

Db 157 GAGCAGAAATACCTAGCTGCTCAGTCACAGAAACATCCAGTTTATGGAGTAATT 216

QY 22 LeuAsnPro-----ValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsn 39

Db 217 CTTGAACCTCCCAACTATGGCATTGTTCACAGAAATATGCTTCTCTGGGATCACTCTATGAT 276

QY 40 ValLeuHisGly-----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSer 57

Db 277 TACATTAAACAGTACAGAAAGTGAGGAGATG-----GATATGGATCACTATATGACC 327

QY 58 TrpCysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeu 77

Db 328 TGGGCCACTGATGTAGCCAAAGGAATGCATTATTTACATATGAGGCTCTCTGCAAGGTG 387

QY 78 IleHisArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLys 97

Db 388 ATTCACAGAGACCTCAAGTCAAGAAACGTTGTATAGCTGCTGATGGA---GTACTGAAG 444

QY 98 IleCysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnLeuLys----- 115

Db 445 ATCTGTGACTTTGGT---GCCTCTCGGTTCCATAACCATACAACACACATGCTCCTGGTT 501

QY 116 GlySerAlaAlaTrpMetAlaProGluValPheGluGlySerAsnTyrSerGlyLysCys 135

Db 502 GGAATTTCCCATGGATGGCTCCAGAGATTATCCAGAGTCTCCCTGTGTCAGAAACTTGT 561

QY 136 AspValPheSerTrpGlyIleLeuTrpGluValIleThrArgArgLysProPheAsp 155

Db 562 GACACATATTCCTATGCTGTGTTCTCTGGAGATGCTAACAGGGAGGTCCTTTTAAA 621

QY 156 GluIleGlyGlyProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgPro 174

Db 622 GGTTTGGAAGGA-----TTACAAGTAGCTTGGCTTGTAGTGAAAAAACAAGAGATT 675

QY 175 ProLeuIleLysAsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLys 194

Db 676 ACCATTCCAGCAGTTCGCCAGAGTTTGTGGAATGTTACATCAGTGTGGGAAGCT 735

QY 195 AspProSerGlnArgProSerMetGluLeuIleValLysIleMet 209

Db 736 GATGCCAAGAAACGCCCATCATTCAGCAAATCATTTCAATCCTG 780

RESULT 7

US-09-221-527-6

; Sequence 6, Application US/09221527

; Patent No. 6146832

; GENERAL INFORMATION:

; APPLICANT: Acton, Susan

; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR

; FILE REFERENCE: WNI-050

; CURRENT APPLICATION NUMBER: US/09/221,527

; CURRENT FILING DATE: 1998-12-28

; EARLIER APPLICATION NUMBER: 09/163,115

; EARLIER FILING DATE:

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 6

; LENGTH: 1365

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(1365)

US-09-221-527-6

Alignment Scores:

Pred. No.: 6.56e-35 Length: 1365
Score: 371.00 Matches: 81
Percent Similarity: 56.28% Conservative: 40
Best Local Similarity: 37.67% Mismatches: 80
Query Match: 29.63% Indels: 14
DB: 3 Gaps: 8

US-09-830-144-2_COPY_76_303 (1-228) x US-09-221-527-6 (1-1365)

QY 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCys 21

Db 157 GAGCAGAAATACCTAGCTGCTCAGTCACAGAAACATCCAGTTTATGGAGTAATT 216

QY 22 LeuAsnPro-----ValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsn 39

Db 217 CTTGAACCTCCCAACTATGGCATTGTTCACAGAAATATGCTTCTCTGGGATCACTCTATGAT 276

QY 40 ValLeuHisGly-----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSer 57

Db 277 TACATTAAACAGTACAGAAAGTGAGGAGATG-----GATATGGATCACTATATGACC 327

QY 58 TrpCysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeu 77

Db 328 TGGGCCACTGATGTAGCCAAAGGAATGCATTATTTACATATGAGGCTCTCTGCAAGGTG 387

QY 78 IleHisArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLys 97

Db 388 ATTCACAGAGACCTCAAGTCAAGAAACGTTGTATAGCTGCTGATGGA---GTACTGAAG 444

QY 98 IleCysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnLeuLys----- 115

Db 445 ATCTGTGACTTTGGT---GCCTCTCGGTTCCATAACCATACAACACACATGCTCCTGGTT 501

QY 116 GlySerAlaAlaTrpMetAlaProGluValPheGluGlySerAsnTyrSerGlyLysCys 135

Db 502 GGAATTTCCCATGGATGGCTCCAGAGTTATCCAGAGTCTCCCTGTGTCAGAAACTTGT 561

QY 136 AspValPheSerTrpGlyIleLeuTrpGluValIleThrArgArgLysProPheAsp 155

Db 562 GACACATATTCCTATGCTGTGTTCTCTGGAGATGCTAACAGGGAGGTCCTTTTAAA 621

QY 156 GluIleGlyGlyProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgPro 174

Db 622 GGTTTGGAAGGA-----TTACAAGTAGCTTGGCTTGTAGTGAAAAAACAAGAGATT 675

QY 175 ProLeuIleLysAsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLys 194

Db 676 ACCATTCCAGCAGTTCGCCAGAGTTTGTGGAATGTTACATCAGTGTGGGAAGCT 735

QY 195 AspProSerGlnArgProSerMetGluLeuIleValLysIleMet 209

Db 736 GATGCCAAGAAACGCCCATCATTCAGCAAATCATTTCAATCCTG 780

RESULT 8

US-09-221-236-6

; Sequence 6, Application US/09221236

; Patent No. 6146841

; GENERAL INFORMATION:

```

; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/221,236
; EARLIER FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1365)
US-09-221-236-6

```

```

Alignment Scores:
Pred. No.: 6,56e-35 Length: 1365
Score: 371.00 Matches: 81
Percent Similarity: 56.28% Conservative: 40
Best Local Similarity: 37.67% Mismatches: 80
Query Match: 29.63% Indels: 14
DB: Gaps: 8

```

US-09-830-144-2_COPY_76_303 (1-228) x US-09-221-236-6 (1-1365)

```

QY 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIysLeuTyrgIyAlaCys 21
DB 157 GAGCGCAAAATACTCTACGTCTCTCAGTCACAGAAACATCATCTCAGTTTATGAGTAATT 216
QY 22 LeuAsnPro-----ValCysLeuValMetGluTyrgIyAlaGluGlySerLeuTyraAsn 39
DB 217 CTGGAACCTCCCAACTAGTGCATTTGCACAGAAATATGCTTCTCTGGAGATCATCTATGAT 276
QY 40 ValLeuHisGly-----AlaGluProLeuProTyrtTyrtThrAlaAlaHisAlaMetSer 57
DB 277 TACATTAACTGTAACAGAAAGTGAAGATG-----GATATGATCATATTATGATGACC 327
QY 58 TrpCysLeuGlnCysSerGlnGlyValAlaIleTyrlleuHisSerMetGlnProIysAlaLeu 77
DB 328 TGGGCCACTGATGATGACCAAGAAATGCAATTATTACATATGAGAGCTCCTGTCAAGGTG 387
QY 78 IleHisArgAspLeuLysProProAsnLeuLeuValAlaGlyIleThrValLeuLys 97
DB 388 ATTCAAGAGAACTCAAGTCAAGAAACGTTTATAGCTGATGGA---GTAAGTGAAG 444
QY 98 IleCysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys----- 115
DB 445 ATCTGTGACTTTGGT---GCCTCTCGGTTCCATTAACCATCAACACACATGTCCTTGTT 501
QY 116 GlySerAlaIleTrpMetAlaProGluValPheGluGlySerAsnTyrsGlnLysCys 135
DB 502 GGAACCTTCCCATGAGAGGCTCCAGAAATATCCAGAGTCTCCCTGTGTCAAAACTTGT 561
QY 136 AspValPheSerTrpGlyIleIleLeuTrpGluValIleThrArgArgLysProPheAsp 155
DB 562 GACACATATTCCTATGCTGTGTCTGTGGAGATGCTAACAGAGAGGTCCTTTTAA 621
QY 156 GluIleGlyIleProAlaPheArgIleMetTrp--AlaValHisAsnGlyThrArgPro 174
DB 622 GGTTTGGAAGA-----TTACAAGTACGTTGGCTTGTAGTGGAAAAAAGAGAGATT 675
QY 175 ProLeuIleLysAsnLeuProLysProIleGluSerIleMetThrArgCysTrpSerLys 194
DB 676 ACCATTCCAAAGCAGTGGCCCAAGAAAGTTTGTGAACTGTATCATCATGTGGGAAGCT 735
QY 195 AspProSerGlnArgProSerMetGluGluIleValIysIleMet 209
DB 736 GATGCCAAGAAAGCGCATATTCACGAACAATCATTTCAATCTCTG 780
RESULT 9

```

```

US-09-221-416-6
; Sequence 6, Application US/09221416
; Patent No. 6153417
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/221,416
; EARLIER FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1365)
US-09-221-416-6

```

```

Alignment Scores:
Pred. No.: 6,56e-35 Length: 1365
Score: 371.00 Matches: 81
Percent Similarity: 56.28% Conservative: 40
Best Local Similarity: 37.67% Mismatches: 80
Query Match: 29.63% Indels: 14
DB: Gaps: 8

```

US-09-830-144-2_COPY_76_303 (1-228) x US-09-221-416-6 (1-1365)

```

QY 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIysLeuTyrgIyAlaCys 21
DB 157 GAGCGCAAAATACTCTACGTCTCTCAGTCACAGAAACATCATCTCAGTTTATGAGTAATT 216
QY 22 LeuAsnPro-----ValCysLeuValMetGluTyrgIyAlaGluGlySerLeuTyraAsn 39
DB 217 CTGGAACCTCCCAACTAGTGCATTTGCACAGAAATATGCTTCTCTGGAGATCATCTATGAT 276
QY 40 ValLeuHisGly-----AlaGluProLeuProTyrtTyrtThrAlaAlaHisAlaMetSer 57
DB 277 TACATTAACTGTAACAGAAAGTGAAGATG-----GATATGATCATATTATGATGACC 327
QY 58 TrpCysLeuGlnCysSerGlnGlyValAlaIleTyrlleuHisSerMetGlnProIysAlaLeu 77
DB 328 TGGGCCACTGATGATGACCAAGAAATGCAATTATTACATATGAGAGCTCCTGTCAAGGTG 387
QY 78 IleHisArgAspLeuLysProProAsnLeuLeuValAlaGlyIleThrValLeuLys 97
DB 388 ATTCAAGAGAACTCAAGTCAAGAAACGTTTATAGCTGATGGA---GTAAGTGAAG 444
QY 98 IleCysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys----- 115
DB 445 ATCTGTGACTTTGGT---GCCTCTCGGTTCCATTAACCATCAACACACATGTCCTTGTT 501
QY 116 GlySerAlaIleTrpMetAlaProGluValPheGluGlySerAsnTyrsGlnLysCys 135
DB 502 GGAACCTTCCCATGAGAGGCTCCAGAAATATCCAGAGTCTCCCTGTGTCAAAACTTGT 561
QY 136 AspValPheSerTrpGlyIleIleLeuTrpGluValIleThrArgArgLysProPheAsp 155
DB 562 GACACATATTCCTATGCTGTGTCTGTGGAGATGCTAACAGAGAGGTCCTTTTAA 621
QY 156 GluIleGlyIleProAlaPheArgIleMetTrp--AlaValHisAsnGlyThrArgPro 174
DB 622 GGTTTGGAAGA-----TTACAAGTACGTTGGCTTGTAGTGGAAAAAAGAGAGATT 675
QY 175 ProLeuIleLysAsnLeuProLysProIleGluSerIleMetThrArgCysTrpSerLys 194
DB 676 ACCATTCCAAAGCAGTGGCCCAAGAAAGTTTGTGAACTGTATCATCATGTGGGAAGCT 735
QY 195 AspProSerGlnArgProSerMetGluGluIleValIysIleMet 209

```


QY 175 ProLeuIIeLysAsnLeuProLysProIIeGluSerLeuMetThrArgCysTrpSerLys 194
Db 676 ACCATTCCAGAGCACTTGGCCCCAGAGTTTGGCTGTAACCTTTACATCAGAGTTGGGAAGCT 735
QY 195 AspProSerGlnArgProSerMetGluGluIleValLysIleMet 209
Db 736 GATGCCAAGAAACGGCCATCATTTCAAGCAAAATCATTTCAATCTCTG 780

RESULT 12
US-09-221-528-6
; Sequence 6, Application US/09221528
; Patent No. 6130874
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,528
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1365)
US-09-221-528-6

Alignment Scores:
Pred. No.: 6,56e-35 Length: 1365
Score: 371.00 Matches: 81
Percent Similarity: 56.28% Conservative: 40
Best Local Similarity: 37.67% Mismatches: 80
Query Match: 29.63% Indels: 14
Gaps: 8

US-09-830-144-2_COPY_76_303 (1-228) x US-09-221-528-6 (1-1365)

QY 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrgIyAlaCys 21
Db 157 GAGGCAAGAAATACTCAGTGTCTCCAGCTCACAGAAATCATCATGTTTATGAGTAATT 216
QY 22 LeuAsnPro-----ValCysLeuValMetGluTyrgIyAlaGluGlyCysLeuTyrgAsn 39
Db 217 CTGGAACCTTCCCAACTATGAGCATTTGTCACAGAAATATGCTTCTGTGGATCAGCTTAATGAT 276
QY 40 ValLeuHisGly-----AlaGluProLeuProTyrgTyrgThrAlaHisAlaMetSer 57
Db 277 TACATTAAACGTAACAGAGTAGAGAGATG-----GATATGAGTACATTATGACC 327
QY 58 TyrCysLeuGlnCysSerGlnGlyValAlaTyrgLeuHisSerMetGlnProLysAlaLeu 77
Db 328 TGGGCCACTGATGTAGCCAAAGAGATGCAATTATTACATATGAGGCTCTCTGTCAAGGTG 387
QY 78 IleHisArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLys 97
Db 388 ATTCACAGAGACCTCAAGTCCAGAAACGTTGTATAGCTGTGAGTGA--GTAAGAAG 444
QY 98 IleCysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnLys----- 115
Db 445 ATCTGTGACTTTGGT---GCCTCGGTTCCATTAACCATTAACAACAACATGCTTTGGTT 501
QY 116 GlySerAlaIleTyrMetAlaProGluValPheGluGlySerAsnTyrgIyLysCys 135
Db 502 GGAACCTTCCCATGATGGCTCCAGAAATTATCCAGAGTCTCCCTGTGCAGAAACTTGT 561
QY 136 AspValPheSerTrpGlyIleIleLeuTrpGluValIleThrArgArgLysProPheAsp 155
Db 562 GACACATATTCATTGATGTGTGTCTGTGGAGATGCTAAACAAGAGGTCCCTTTAA 621

QY 156 GluIleGlyGlyProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgPro 174
Db 622 GGTTTGGAGAAGA-----TTACAAGTAGCTTGCTGTGATGGAAAAACAGAGATTAA 675
QY 175 ProLeuIIeLysAsnLeuProLysProIIeGluSerLeuMetThrArgCysTrpSerLys 194
Db 676 ACCATTCCAGAGCACTTGGCCCCAGAGTTTGGCTGTAACCTTTACATCAGAGTTGGGAAGCT 735
QY 195 AspProSerGlnArgProSerMetGluGluIleValLysIleMet 209
Db 736 GATGCCAAGAAACGGCCATCATTTCAAGCAAAATCATTTCAATCTCTG 780

RESULT 13
US-09-593-553-6
; Sequence 6, Application US/09593553
; Patent No. 6200770
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/593,553
; CURRENT FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: 09/163,115
; PRIOR FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1365)
US-09-593-553-6

Alignment Scores:
Pred. No.: 6,56e-35 Length: 1365
Score: 371.00 Matches: 81
Percent Similarity: 56.28% Conservative: 40
Best Local Similarity: 37.67% Mismatches: 80
Query Match: 29.63% Indels: 14
Gaps: 8

US-09-830-144-2_COPY_76_303 (1-228) x US-09-593-553-6 (1-1365)

QY 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrgIyAlaCys 21
Db 157 GAGGCAAGAAATACTCAGTGTCTCCAGCTCACAGAAACATCTCCAGTTTATGAGTAATT 216
QY 22 LeuAsnPro-----ValCysLeuValMetGluTyrgIyAlaGluGlyCysLeuTyrgAsn 39
Db 217 CTGGAACCTTCCCAACTATGAGCATTTGTCACAGAAATATGCTTCTGTGGATCAGCTTAATGAT 276
QY 40 ValLeuHisGly-----AlaGluProLeuProTyrgTyrgThrAlaHisAlaMetSer 57
Db 277 TACATTAAACGTAACAGAGTAGAGAGATG-----GATATGAGTACATTATGACC 327
QY 58 TyrCysLeuGlnCysSerGlnGlyValAlaTyrgLeuHisSerMetGlnProLysAlaLeu 77
Db 328 TGGGCCACTGATGTAGCCAAAGAGATGCAATTATTACATATGAGGCTCTCTGTCAAGGTG 387
QY 78 IleHisArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLys 97
Db 388 ATTCACAGAGACCTCAAGTCCAGAAACGTTGTATAGCTGTGAGTGA--GTAAGAAG 444
QY 98 IleCysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnLys----- 115
Db 445 ATCTGTGACTTTGGT---GCCTCGGTTCCATTAACCATTAACAACAACATGCTTTGGTT 501
QY 116 GlySerAlaIleTyrMetAlaProGluValPheGluGlySerAsnTyrgIyLysCys 135
Db 502 GGAACCTTCCCATGATGGCTCCAGAAATTATCCAGAGTCTCCCTGTGCAGAAACTTGT 561

```
Qy 136 AspValPheSerTrpGlyIleLeuTrpGluValIleThrArgArgLysProPheAsp 155
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 562 GACACATATCTCTATGGTGTGGTCTCTGGGAGATGCTAACAGGGAGGTCCTCTTTAAA 621

Qy 156 GluIleGlyGlyProAlaPheArgIleMetTrp---AlaValHisnGlyThrArgPro 174
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 622 GGTITGGNAGA-----TTACAAGTAGCTTGGCTTGTAGTGGAAAAAACAAGAGATTA 675

Qy 175 ProLeuIleLysAsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLys 194
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 676 ACCATTCCAGCAGTTGCCCCAGAAGTTTTCGTAACACTGTTACATCAGTGTGGGAAGCT 735

Qy 195 AspProSerGlnArgProSerMetGluGluIleValLysIleMet 209
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 736 GATGCCAAGAAACGCCCATCATTTCAAGCAAAATCATTTCAATCCTG 780

RESULT 14
US-09-221-237-6
; Sequence 6, Application US/09221237
; Patent No. 6214597
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,237
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1365)
US-09-221-237-6

Alignment Scores:
Pred. No.: 6 56e-35 Length: 1365
Score: 371.00 Matches: 81
Percent Similarity: 56.28% Conservative: 40
Best Local Similarity: 37.67% Mismatches: 80
Query Match: 29.63% Indels: 14
DB: 4 Gaps: 6

US-09-830-144-2_COPY_76_303 (1-228) x US-09-221-237-6 (1-1365)

Qy 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrglyAlaCys 21
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 157 GAGCAGAAATACTCAGTGCTCCTCAGTCACAGAAACATCATCCAGTTTATGGAGTAATT 216

Qy 22 LeuAsnPro-----ValCysLeuValMetGluTyraAlaGluGlyGlySerLeuTyraAsn 39
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 217 CTTGAACCTCCCAACTATGTCATTGTTCACAGAATATGCTTCTCTGGGATCACTCTATGAT 276

Qy 40 ValLeuHisGly-----AlaGluProLeuProTyTyThrAlaAlaHisAlaMetSer 57
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 277 TACATTAACTAGTACAGAAAGTGGAGATG-----GATATGATCACAATATGACC 327

Qy 58 TrpCysLeuGlnCysSerGlnGlyValAlaTyLeuHisSerMetGlnProLysAlaLeu 77
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 328 TGGGCCACTGATGAGCAAGGAATGCATTATTACATATGAGGCTCCTGTCAAGGTG 387

Qy 78 IleHisArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLys 97
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 388 ATTCAAGAGACCTCAAGTCAAGAAACGTTGTATAGTCTGATGGA---GTACTGAAG 444

Qy 98 IleCysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys----- 115
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 445 ATCTGTGACTTTGGT---GCCTCTCGGTTCCATAACACATCAACACACATGCTTGGTT 501
```

```
Qy 116 GlySerAlaAlaTrpMetAlaProGluValPheGluGlySerAsnTySerGluLysCys 135
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 502 GGAACCTTTCCCATGGATGGCTCCAGAAGTTATCCAGAGTCTCCCTGTGTACAGAACTTGT 561

Qy 136 AspValPheSerTrpGlyIleLeuTrpGluValIleThrArgArgLysProPheAsp 155
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 562 GACACATATCTCTATGGTGTGGTCTCTGGGAGATGCTAACAGGGAGGTCCTCTTTAAA 621

Qy 156 GluIleGlyGlyProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgPro 174
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 622 GGTITGGNAGA-----TTACAAGTAGCTTGGCTTGTAGTGGAAAAAACAAGAGATTA 675

Qy 175 ProLeuIleLysAsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLys 194
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 676 ACCATTCCAGCAGTTGCCCCAGAAGTTTTCGTAACACTGTTACATCAGTGTGGGAAGCT 735

Qy 195 AspProSerGlnArgProSerMetGluGluIleValLysIleMet 209
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 736 GATGCCAAGAAACGCCCATCATTTCAAGCAAAATCATTTCAATCCTG 780

RESULT 15
US-09-221-235-4
; Sequence 4, Application US/09221235
; Patent No. 6043040
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,235
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2120
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (47)..(1411)
US-09-221-235-4

Alignment Scores:
Pred. No.: 1 27e-34 Length: 2120
Score: 371.00 Matches: 81
Percent Similarity: 56.28% Conservative: 40
Best Local Similarity: 37.67% Mismatches: 80
Query Match: 29.63% Indels: 14
DB: 3 Gaps: 6

US-09-830-144-2_COPY_76_303 (1-228) x US-09-221-235-4 (1-2120)

Qy 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrglyAlaCys 21
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 203 GAGCAGAAATACTCAGTGCTCCTCAGTCACAGAAACATCATCCAGTTTATGGAGTAATT 262

Qy 22 LeuAsnPro-----ValCysLeuValMetGluTyraAlaGluGlyGlySerLeuTyraAsn 39
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 263 CTTGAACCTCCCAACTATGTCATTGTTCACAGAATATGCTTCTCTGGGATCACTCTATGAT 322

Qy 40 ValLeuHisGly-----AlaGluProLeuProTyTyThrAlaAlaHisAlaMetSer 57
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 323 TACATTAACTAGTACAGAAAGTGGAGATG-----GATATGATCACAATATGACC 373

Qy 58 TrpCysLeuGlnCysSerGlnGlyValAlaTyLeuHisSerMetGlnProLysAlaLeu 77
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 374 TGGGCCACTGATGAGCAAGGAATGCATTATTACATATGAGGCTCCTGTCAAGGTG 433

Qy 78 IleHisArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLys 97
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 434 ATTCACAGAGACCTCAAGTCAAGAAACGTTGTATAGTCTGATGGA---GTACTGAAG 490
```

```

Qy 98 IleCysaspPheGlyThrAlaCysaspIleGlnThrHisMetThrAsnAsnLys----- 115
Db 491 ATCTGTGACTTTGGT---GCCCTCGGTTCCATAACCATACAAACACATGTCCTTGGTT 547
Qy 116 GlySerAlaAlaTrpMetAlaProGluValPheGluGlySerAsnTy-SerGluLysCys 135
Db 548 GGAACCTTCCAGATGGCTCCAGAAAGTTATCCAGAGTCTCCCTGTGTCAAGAACTTGT 607
Qy 136 AspValPheSerTrpGlyIleIleLeuTrpGluValIleThrArgArgLysProPheAsp 155
Db 608 GACACATATTCTTANGTGTGTGCTCTGTGGAGATGCTACAGAGGAGGTCCCTTTAA 667
Qy 156 GluIleGlyIleProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgPro 174
Db 668 GGTTCGAAGA---TTACAGTAGCTTGCTGTAGTGAAAAAACAAGAGATT 721
Qy 175 ProLeuIleLysAsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLys 194
Db 722 ACCATTCCAAAGCAGTTGCCCCAGAAAGTTTGCTGAACGTGTACATCAGTGTGGAGACT 781
Qy 195 AspProSerGlnArgProSerMetGluGluIleValLysIleMet 209
Db 782 GATGCCAAGAAAGGCCATCATTTCAAGCAAAATCATTTCAATCCTG 826

```

Search completed: December 10, 2002, 05:36:02
 Job time : 74 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 10, 2002, 03:59:56 ; Search time 913 Seconds

(without alignments)
97.376 Million cell updates/sec

Title: US-09-830-144-2_COPY_76_303

Perfect score: 1252

Sequence: 1 VELRQLSRVNHFNIVKLXGA.....MTHLMRYPPGADPELOYPQC 228

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 350425 seqs, 194966369 residues

Total number of hits satisfying chosen parameters: 700850

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlp
-Q/cgn2_1/USPTO_spool/US09830144/runat_04122002_141755_4398/app_query.fasta_1.391
-DB=Published Applications NA -QWMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09830144 @CGN 1.1 36 @runat_04122002_141755_4398
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

```
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1252	100.0	1788	9	US-10-158-895-14
2	1252	100.0	2856	9	US-10-158-895-3
3	392.5	31.3	3454	10	US-09-969-347-226
4	371	29.6	1365	10	US-09-757-982-6

ALIGNMENTS

RESULT 1

US-10-158-895-14

; Sequence 14, Application US/10158895

; Patent No. US20020155624A1

; GENERAL INFORMATION:

; APPLICANT: ONO, KOICHIRO

; APPLICANT: ONTOMO, TOSHIHIKO

; APPLICANT: TSUCHIYA, MASAYUKI

; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES

; FILE REFERENCE: 053466/0278

; CURRENT APPLICATION NUMBER: US/10/158,895

; CURRENT FILING DATE: 2002-06-03

; PRIOR APPLICATION NUMBER: US/09/529,279

; PRIOR FILING DATE: 2000-04-11

; PRIOR APPLICATION NUMBER: PCT/JP98/04796

; PRIOR FILING DATE: 1998-10-22

; PRIOR APPLICATION NUMBER: JP 9/290188

; PRIOR FILING DATE: 1997-10-22

; NUMBER OF SEQ ID NOS: 48

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 14

; LENGTH: 1788

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (7)..(1776)

US-10-158-895-14

Sequence 4, Appli
Sequence 1, Appli
Sequence 3, Appli
Sequence 882, App
Sequence 1014, Ap
Sequence 3, Appli
Sequence 1, Appli
Sequence 9, Appli
Sequence 7, Appli
Sequence 903, App
Sequence 1577, Ap
Sequence 1, Appli
Sequence 1073, Ap
Sequence 84, Appl
Sequence 405, Appl
Sequence 2, Appli
Sequence 16, Appl
Sequence 1479, Ap
Sequence 1085, Ap
Sequence 48, Appl
Sequence 271, App
Sequence 1405, Ap
Sequence 1611, Ap
Sequence 82, Appl
Sequence 1, Appli
Sequence 1, Appli
Sequence 553, App
Sequence 1, Appli
Sequence 2, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 5, Appli
Sequence 293, App
Sequence 1599, Ap
Sequence 1, Appli
Sequence 1983, Ap
Sequence 250, App
Sequence 1, Appli
Sequence 13, Appl
Sequence 7, Appli

10 US-09-757-982-4
12 US-10-014-882-1
12 US-10-014-882-3
9 US-09-938-842A-882
9 US-09-938-842A-1014
10 US-09-947-199-3
10 US-09-947-199-9
10 US-09-947-199-9
10 US-09-947-199-7
9 US-09-938-842A-903
9 US-09-938-842A-1577
10 US-09-904-389-1
9 US-09-938-842A-1073
10 US-09-919-172-84
9 US-09-954-531-405
10 US-09-735-103-2
9 US-09-938-842A-2
12 US-10-045-428A-2
10 US-09-828-313-16
9 US-09-938-842A-1479
9 US-09-938-842A-1085
12 US-10-044-090-48
10 US-09-925-302-271
10 US-09-917-800A-1405
10 US-09-917-800A-1611
10 US-09-728-952-82
10 US-09-977-269-1
10 US-09-731-231A-1
9 US-09-938-842A-553
10 US-09-757-100B-1
10 US-09-906-397-2
10 US-09-758-003-1
10 US-09-906-397-1
10 US-09-906-397-5
10 US-09-954-456-293
10 US-09-954-456-1599
12 US-10-000-864-1
10 US-09-954-456-1983
9 US-10-108-605-250
10 US-09-858-754-1
10 US-09-858-754-13
12 US-10-000-864-7

C

371

29.6

2120

10

US-09-757-982-4

371

29.6

2120

10

US-09-757-982-4

355

28.4

3111

12

US-10-014-882-1

3518

28.4

3518

12

US-10-014-882-3

347

27.7

1428

9

US-09-938-842A-882

345.5

27.6

1662

9

US-09-938-842A-1014

338.5

27.0

2505

10

US-09-947-199-3

338.5

27.0

3025

10

US-09-947-199-9

335.5

26.8

2505

10

US-09-947-199-9

335.5

26.8

3026

10

US-09-947-199-7

329.5

26.3

1638

9

US-09-938-842A-903

328.5

26.2

2211

9

US-09-938-842A-1577

313.5

25.0

3886

10

US-09-904-389-1

304.5

24.3

2892

9

US-09-938-842A-1073

301.5

24.1

3663

9

US-09-919-172-84

298.5

23.8

6383

9

US-09-954-531-405

233.5

23.4

2574

10

US-09-735-103-2

293.5

23.4

2574

12

US-10-045-428A-2

289

23.1

1627

10

US-09-828-313-16

288

23.0

3012

9

US-09-938-842A-1479

285

22.8

2661

9

US-09-938-842A-1085

278

22.2

3454

12

US-10-044-090-48

278

22.2

3726

10

US-09-925-302-271

271.5

21.7

2719

10

US-09-917-800A-1405

270.5

21.6

1911

10

US-09-917-800A-1611

270

21.6

2184

10

US-09-728-952-82

269.5

21.5

2000

10

US-09-977-269-1

268.5

21.4

2869

10

US-09-731-231A-1

268

21.4

1052

9

US-09-938-842A-553

268

21.4

3791

10

US-09-757-100B-1

266.5

21.3

1296

10

US-09-906-397-2

266.5

21.3

2016

10

US-09-758-003-1

266.5

21.3

2322

10

US-09-906-397-1

266.5

21.3

2322

10

Alignment Scores:

Pred. No.: 1,33e-138 Length: 1788
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-830-144-2_copy_76_303 (1-228) x US-10-158-895-14 (1-1788)

```
QY 1 ValGluLeuArgGluLeuSerArgValAlaSnHISProAsnIleValIysLeuTYrGlyAla 20
    |||
Db 232 GTAGAGCTTCGGAGCTTATCCGCTGGAACCACTCTTAATATTGTAAGCTTTATGAGACC 291

QY 21 CysLeuAsnProValCysLeuValMetGluTYrAlaGluGlyIseLeuTYrAsnVal 40
    |||
Db 292 TGCTTGAATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGGCTCTTATATATGTG 351

QY 41 LeuHISGlyAlaGluProLeuProTYrTYrThrAlaAlaHISAlaMetSerTrpCysLeu 60
    |||
Db 352 CTGCATGCTGCTGACCACTTGCATATTACTGCTGCCACGCAATGATGAGTGTGTTTA 411

QY 61 GluCysSerGluGlyValAlaTYrLeuHISSerMetGluProIysAlaLeuIleHISArg 80
    |||
Db 412 CAGTGTTCCTCCAGAGAGCTGCTTATCTTCAAGCATGCAACCCAAAGCGCTTAATTCACAG 471

QY 81 AspLeuIysProProAsnLeuLeuValAlaGlyGlyThrValLeuIysIleCysAsp 100
    |||
Db 472 GACCTGAACCAACCAACTTACTGCTGTGTCAGGGGGGACGTTCTAATAATTTGTGAT 531

QY 101 PheGlyThrAlaCysAspIleGlnThrHISMetThrAsnAsnIysGlySerAlaAlaTrp 120
    |||
Db 532 TTGGGTACAGCCCTGTGACATTCAGACACATGACCAATTAACAGGGGAGTGTCTGTGG 591

QY 121 MetAlaProGluValPheGluGlySerAsnTYrSerGluIysCysAspValPheSerTrp 140
    |||
Db 592 ATGGCACCTGAAAGTTTGTGAAGGTAGTAATTACAGTGAATAAATGTGACCTTCCACGTGG 651

QY 141 GlyIleIleLeuTrpGluValIleThrArgArgIysProPheAspGluIleGlyIlePro 160
    |||
Db 652 GGTATTATTCTTTGGGAGAGATGATACGGCTCGGAACCTTTGATGAGATTGTGGCCCA 711

QY 161 AlaPheArgIleMetTrpAlaValHISAsnGlyThrArgProProLeuIleIysAsnLeu 180
    |||
Db 712 GCTTCCGATCATGTGGGCTGTTCATAATGTACTCGACCACTGATTAATAAATTTTA 771

QY 181 ProIysProIleGluSerLeuMetThrArgCysTrpSerIysAspProSerGluArgPro 200
    |||
Db 772 CCTAAGCCCATTTAGAGCTGATGACTCGTTGTTGGTCTTAAGATCTTCCACGGCCCT 831

QY 201 SerMetGluGluIleValIysIleMetThrHISLeuMetArgTYrPheProGlyAlaAsp 220
    |||
Db 832 TCATGAGGAAATTGTGAATAATATGACCTCATGTGGGGTACTTCCAGAGAGAGAT 891

QY 221 GluProLeuGlnTYrProCysGln 228
    |||
Db 892 GAGCCATTACAGTATCTTGTCTAG 915

RESULT 2
US-10-158-895-3
; Sequence 3, Application US/10158895
; Patent No. US20020155624A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF- $\beta$  INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
```

PRIOR APPLICATION NUMBER: PCT/JP98/04796

PRIOR FILING DATE: 1998-10-22

PRIOR APPLICATION NUMBER: JP 9/290188

PRIOR FILING DATE: 1997-10-22

NUMBER OF SEQ ID NOS: 48

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 3

LENGTH: 2656

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (183)..(1919)

US-10-158-895-3

Alignment Scores:

Pred. No.: 2.37e-138 Length: 2656
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-830-144-2_copy_76_303 (1-228) x US-10-158-895-3 (1-2656)

```
QY 1 ValGluLeuArgGluLeuSerArgValAlaSnHISProAsnIleValIysLeuTYrGlyAla 20
    |||
Db 408 GTAGAGCTTCGGAGCTTATCCGCTGGAACCACTCTTAATATTGTAAGCTTTATGAGACC 467

QY 21 CysLeuAsnProValCysLeuValMetGluTYrAlaGluGlyIseLeuTYrAsnVal 40
    |||
Db 468 TGCTTGAATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGGCTCTTATATATGTG 527

QY 41 LeuHISGlyAlaGluProLeuProTYrTYrThrAlaAlaHISAlaMetSerTrpCysLeu 60
    |||
Db 528 CTGCATGCTGCTGACCACTTGCATATTACTGCTGCCACGCAATGATGAGTGTGTTTA 587

QY 61 GluCysSerGluGlyValAlaTYrLeuHISSerMetGluProIysAlaLeuIleHISArg 80
    |||
Db 588 CAGTGTTCCTCCAGAGAGCTGCTTATCTTCAAGCATGCAACCCAAAGCGCTTAATTCACAG 647

QY 81 AspLeuIysProProAsnLeuLeuValAlaGlyGlyThrValLeuIysIleCysAsp 100
    |||
Db 648 GACCTGAACCAACCAACTTACTGCTGTGTCAGGGGGGACGTTCTAATAATTTGTGAT 707

QY 101 PheGlyThrAlaCysAspIleGlnThrHISMetThrAsnAsnIysGlySerAlaAlaTrp 120
    |||
Db 708 TTGGGTACAGCCCTGTGACATTCAGACACATGACCAATTAACAGGGGAGTGTCTGTGG 767

QY 121 MetAlaProGluValPheGluGlySerAsnTYrSerGluIysCysAspValPheSerTrp 140
    |||
Db 768 ATGGCACCTGAAAGTTTGTGAAGGTAGTAATTACAGTGAATAAATGTGACGTTCACTGG 827

QY 141 GlyIleIleLeuTrpGluValIleThrArgArgIysProPheAspGluIleGlyIlePro 160
    |||
Db 828 GGTATTATTCTTTGGGAGAGATGATACGGCTCGGAACCTTTGATGAGATTGTGGCCCA 887

QY 161 AlaPheArgIleMetTrpAlaValHISAsnGlyThrArgProProLeuIleIysAsnLeu 180
    |||
Db 888 GCTTCCGATCATGTGGGCTGTTCATAATGTACTCGACCACTGATTAATAAATTTTA 947

QY 181 ProIysProIleGluSerLeuMetThrArgCysTrpSerIysAspProSerGluArgPro 200
    |||
Db 948 CCTAAGCCCATTTAGAGCTGATGACTCGTTGTTGGTCTTAAGATCTTCCACGGCCCT 1007

QY 201 SerMetGluGluIleValIysIleMetThrHISLeuMetArgTYrPheProGlyAlaAsp 220
    |||
Db 1008 TCATGAGGAAATTGTGAATAATATGACCTCATGTGGGGTACTTCCAGAGAGAGAT 1067

QY 221 GluProLeuGlnTYrProCysGln 228
    |||
Db 1068 GAGCCATTACAGTATCTTGTCTAG 1091
```

RESULT 3

US-09-969-347-226
; Sequence 226, Application US/09969347
; Patent No. US20020115085A1
; GENERAL INFORMATION:

; APPLICANT: Ebner, Reinhard
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu

; FILE OF INVENTION: Sets
; FILE REFERENCE: 689290-69
; CURRENT APPLICATION NUMBER: US/09/969,347

; CURRENT FILING DATE: 2001-10-02

; PRIOR APPLICATION NUMBER: US/60/237,598

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: US/60/237,604

; PRIOR FILING DATE: 2000-10-03

; NUMBER OF SEQ ID NOS: 318

; SOFTWARE: Patentin version 3.0

; SEQ ID NO 226

; LENGTH: 3454

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-969-347-226

Alignment Scores:

Pred. No.: 1.7e-36 Length: 3454
Score: 392.50 Matches: 88
Percent Similarity: 56.68% Conservative: 35
Best Local Similarity: 40.55% Mismatches: 77
Query Match: 31.35% Indels: 17
DB: 10 Gaps: 6

US-09-830-144-2_COPY_76_303 (1-228) x US-09-969-347-226 (1-3454)

Qy 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIysLeuTyrGlyAlaCys 21
Db 721 GAAGCCGGCTCTTTGAGCGCCCTGACAGCCCAACCAATATGCCCTTAGGGCGGCTGC 780
Qy 22 LeuAsnPro-----ValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsn 39
Db 781 CTCAACCCCCACACCTCTCGCTAGTAGTGAGTATGCCGGGGTGTGCACCTGACGAGG 840
Qy 40 ValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCys 59
Db 841 GTGCTGCGAGGTGCGCGGGTGCCACCTCACGTG-----CTGGTCAACTGGGCT 888
Qy 60 LeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHis 79
Db 889 GTGAGGTGCGCGGGCATGAACCTACCTACCAATGATGCCCTCTGCCCATCATCCAC 948
Qy 80 ArgAspLeuLysProProAsnLeuLeuVal-----AlaGly 92
Db 949 CGGACCTCAAGTCCATCAACATCTGATCTGGAGGCCATCGAAGCCACCAACCTCGCA 1008
Qy 93 GlyThrValLeuLysIleCysAspPheGlyThrAlaCysAsp---IleGlnThrHisMet 111
Db 1009 GACACGGTGTCAAGATCAAGGACTTCGGGCTCGCCCGAGTGGGCACCAAGACCAAG 1068
Qy 112 ThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGluValPheGluGlySerAsnTyr 131
Db 1069 ATGAGCGCTCGGGGACCTACGCGCTGATGCGCGCGAGGTTATCCGTCTCTCCCTCTTC 1128
Qy 132 SerGluLysCysAspValPheSerTrpGlyIleIleLeuTrpGluValIleThrArgArg 151
Db 1129 TCCAAAGACGAGTGTCTCGAGCTTCGGGGTGTGCTGTGTGGAGCTGTGACGGGGAG 1188
Qy 152 LysProPheAspGluIleGlyProAlaPheArgIleMetTrpAlaVal---HisAsn 170
Db 1189 GTCCCTACCGTGAGATC-----GACGCCCTTGGCGCGGTATGGCGTGGGTATGAAT 1242
Qy 171 GlyThrArgProProLeuIleLysAsnLeuProLysProIleGluSerLeuMetThrArg 190
Db 1243 AAGCTGACGCTGCCCATTCCTCCACGTGCGCCGAGCCCTTTGCCCGCCTCTCTGGAGAA 1302

Qy 191 CysTrpSerLysAspProSerGlnArgProSerMetGluGluIleValLys 207
Db 1303 TGCTGGGACCCAGACCCCGGCGGCCAGATTTCGGTAGCATCTTTGAAG 1353

RESULT 4

US-09-757-982-6

; Sequence 6, Application US/09757982

; Patent No. US20020094559A1

; GENERAL INFORMATION:

; APPLICANT: Acton, Susan

; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR

; FILE REFERENCE: MNI-050

; CURRENT APPLICATION NUMBER: US/09/757,982

; CURRENT FILING DATE: 2001-01-10

; PRIOR APPLICATION NUMBER: 09/163,115

; PRIOR FILING DATE: 1998-09-29

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 6

; LENGTH: 1365

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(1365)

US-09-757-982-6

Alignment Scores:

Pred. No.: 1.54e-34 Length: 1365
Score: 371.00 Matches: 81
Percent Similarity: 56.28% Conservative: 40
Best Local Similarity: 37.67% Mismatches: 80
Query Match: 29.63% Indels: 14
DB: 10 Gaps: 8

US-09-830-144-2_COPY_76_303 (1-228) x US-09-757-982-6 (1-1365)

Qy 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIysLeuTyrGlyAlaCys 21
Db 157 GAGCAGAAATACTCAGTGTCTCAGTCAGACAGAAACATCATCCAGTTTATGGAGTAATT 216
Qy 22 LeuAsnPro-----ValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsn 39
Db 217 CTTGAACCTCCCAACTATGGCAATGTGCAGAAATATGCTTCTCTGGGATCACTCTATGAT 276
Qy 40 ValLeuHisGly-----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSer 57
Db 277 TACATTAACAGTAACAGAAAGTGAGGAGATG-----GATATGGATCACATTTATGACC 327
Qy 58 TrpCysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeu 77
Db 328 TGGCCCACTGATGTAGCCAAAGGAATGCATTTATACATATGGAGGCTCCTGTCAAGGTG 387
Qy 78 IleHisArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLys 97
Db 388 ATTCACAGAGACCTCAAGTCAAGAAACGTTTATAGTGTGATGGA--GTACTGAAG 444
Qy 98 IleCysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys----- 115
Db 445 ATCTGTGACTTTGGT--GCCTCTCGGTTCCATAACCATACACACATGTCCTTGGTT 501
Qy 116 GlySerAlaAlaTrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCys 135
Db 502 GGAATTTCCCATGGATGGCTCCAGAAAGTATATCCAGAGTCTCCCTGTGTGAGAACTTGT 561
Qy 136 AspValPheSerTrpGlyIleIleLeuTrpGluValIleThrArgArgLysProPheAsp 155
Db 562 GACACATATTCCTATGGTGTGTTCTCTGGAGATGCTAAACAAGGAGGTCCTCCCTTTAAA 621
Qy 156 GluIleGlyGlyProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgPro 174
Db 622 GGTTTTGAAGGA-----TTACAAAGTAGCTTGGCTTTAGTGGAAAAAACAACGAGAGATTA 675

Qy 175 Proleu1leuysanleuprolysPro1leGluserleuMetThrArgCysTrpSerLys 194
Db 676 ACCATTCCAGCAGGTCGCCCAAGAGTTTGTCTGAACCTGTATCATGAGTGGGAAGCT 735
Qy 195 AspProSerGlnArgProSerMetGluGlu1leValLys1leMet 209
Db 736 GATGCCAAGAAAGGCCATCATTCACAGCAATCATTTCAATCTGT 780

RESULT 5

US-09-757-982-4
; Sequence 4, Application US/09757982
; Patent No. US20020094559A1
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/757,982
; CURRENT FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 09/163,115
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2120
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (47)..(1411)
US-09-757-982-4

Alignment Scores:
Pred. No.: 2,92e-34 Length: 2120
Score: 371.00 Matches: 81
Percent Similarity: 56.28% Conservative: 40
Best Local Similarity: 37.67% Mismatches: 80
Query Match: 29.63% Indels: 14
DB: 10 Gaps: 8

US-09-830-144-2_COPY_76_303 (1-228) x US-09-757-982-4 (1-2120)

Qy 2 GluLeuArgGlnLeuSerArgValAsnHisProAsn1leValLysleuTrg1yAlaCys 21
Db 203 GAGGCGAATAATCTCAGTCTCCTCAGTCAACAGAAATCATCAGTTTATGAGTAATT 262
Qy 22 LeuAsnPro-----ValCysleuValMetGluTrgAlaGlu1yG1ySerleuTrgAsn 39
Db 263 CTTGAACTCCCAACTAGTATGAGTGTGCACAGAAATATGCTTCTGGGATCATCTATGAT 322
Qy 40 ValLeuHisG1y-----AlaGluProLeuProTrgTrgThraAlaHisAlaMetSer 57
Db 323 TACATTAAACAGTAACAGAGTGGAGATG-----GATATGAGTACATTATATGACC 373
Qy 58 TrpCysleuGlnCysSerGlnG1yValAlaTrgLeuHisSerMetGlnProLysAlaLeu 77
Db 374 TGGGCCCATGATGTACCAAGGAATGATCTATTATACATATGAGAGCTCTCTCAAGGTG 433
Qy 78 1leHisArgAspLeuLysProProAsnLeuLeuValAlaG1yG1yThrValLeuLys 97
Db 434 ATTCAAGAGAGCTCAAGTCAAGAAACGTTGTTATAGCTGCTGATGGA---CTACTGAAG 490
Qy 98 1leCysAspPheG1yThra1yCysAsp1leGlnThHisMetThrAsnAsnLys----- 115
Db 491 ATCTGTGACTTGGT---CCCTCTCGGTTCCATTAACTATCAACACATGTCCTTGGTT 547
Qy 116 GlySerAlaAlaTrpMetAlaProGluValAlaPheGluG1ySerAsnTrpSerGlnLysCys 135
Db 548 GGAACCTTCCCATGATGAGCTCCAGAGTTATCCAGAGTCTCCCTGTGCAGAAACTTGT 607
Qy 136 AspValPheSerTrpG1yTrg1y1le1leLeuTrpGluValAla1leThrArgArgLysProPheAsp 155
Db 608 GACACATATTCATGTGTGTGTCTCTGTGGAGATGCTTAACAAGGAGGTCCCTTTAA 667

Qy 156 Glu1leG1yG1yProAlaPheArg1leMetTrp---AlaValHisAsnG1yThraArgPro 174
Db 668 GGTGGAGAGA-----TTACAGATGCTTGGCTGTGATGAGAAAAAAGAGAGATT 721
Qy 175 Proleu1leuysanleuprolysPro1leGluserleuMetThrArgCysTrpSerLys 194
Db 722 ACCATTCCAGCAGGTCGCCCAAGAGTTTGTCTGAACCTGTATCATGAGTGGGAAGCT 781

Qy 195 AspProSerGlnArgProSerMetGluGlu1leValLys1leMet 209
Db 782 GATGCCAAGAAAGGCCATCATTCACAGCAATCATTTCAATCTGT 826

RESULT 6

US-10-014-882-1
; Sequence 1, Application US/10014882
; Patent No. US20020107384A1
; GENERAL INFORMATION:
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James
; APPLICANT: Donoho, Gregory
; TITLE OF INVENTION: No. US20020107384A1 Human Kinase and Polynucleotides Encoding t
; FILE REFERENCE: LEX-0279-USA
; CURRENT APPLICATION NUMBER: US/10/014,882
; CURRENT FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: US 60/254,744
; PRIOR FILING DATE: 2000-12-11
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3111
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-014-882-1

Alignment Scores:
Pred. No.: 3,99e-32 Length: 3111
Score: 355.00 Matches: 83
Percent Similarity: 53.48% Conservative: 40
Best Local Similarity: 36.09% Mismatches: 85
Query Match: 28.35% Indels: 22
DB: 12 Gaps: 6

US-09-830-144-2_COPY_76_303 (1-228) x US-10-014-882-1 (1-3111)

Qy 2 GluLeuArgGlnLeuSerArgValAsnHisProAsn1leValLysleuTrg1yAlaCys 21
Db 511 GAGGCTCGGCTCTTCCGATGCTGCGGACCCCAACATCATCGAGTGGCGGCGGTG 570
Qy 22 LeuAsnPro-----ValCysleuValMetGluTrgAlaGlu1yG1ySerleuTrgAsn 39
Db 571 CTGCAGAGCGGACCTCTCTGCTGTGAGATTGCGCGCGGAGGCGGCAACCGA 630
Qy 40 ValLeuHisG1yAlaGluProLeuProTrgTrgThraAla----- 52
Db 631 GCGGTGCGCGCTGCAACGCCGCCCGGAGCCCGCGCGCGCGCGCGCGCGCG 690
Qy 53 -----AlaHisAlaMet---SerTrpCysleuGlnCysSerGlnG1yValAlaTrg 68
Db 691 CGCATCCCTCGACCGTGTGATCACTGCGCGCGCGCGCGCGCGCGCGCGCGCTTC 750
Qy 69 LeuHisSerMetGlnProLysAlaLeu1leHisArgAspLeuLysProProAsnLeuLeu 88
Db 751 CTGCATAGAGAGGCTTGTGCTCCATCTGCAACCGGAGCTCAAGTCCAGCAACATTTG 810
Qy 89 LeuValAlaG1y-----G1yThrValLeuLys1leCysAspPhe 101
Db 811 CTACTTGAAGATAGAACTGATGACATGTGCATTAATAACTTTGAAAGTTACAGATT 870
Qy 102 G1yThra1yCysAsp---1leGlnThHisMetThrAsnAsnLysG1ySerAlaAlaTrp 120
Db 871 GGGTTGGCGAGGAGATGGCAGAGACACCAAAATAGACACAGACAGCTATGCTGG 930
Qy 121 MetAlaProGluValPheGluG1ySerAsnTrpSerGlnLysCysAspValPheSerTrp 140


```
Db 748 CTGCACAGCTCGAGCAGAAATCCCTTCT-----TTGGAACAGCTATCATTTT 798
QY 59 CysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProlysAlaLeuIle 78
Db 799 GGCTCTGGATATTGCTTAAGGAATGGAATATATTCACTCA-----AGAGAGATAGTT 849
QY 79 HisAspAspLeuLysProProHisLeuLeuValAlaGlyGlyThrValLeuLysIle 98
Db 850 CATCGAGATCTGAAGCCCAAAAAC---GTGTGATCGACATGACTTTCACCTGGAAGATT 906
QY 99 CysAspPheGlyThrAlaCysAspIleGln-----ThriMetThrAsnAsnLysGly 116
Db 907 GCTGACTTTGGCATAGCCGTCGAGAGAGTACTGTGATGTTTGGGGGATTAACATAAGA 966
QY 117 SerAlaAlaTyrMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAsp 136
Db 967 ACTTATAGCTGATGGCAGCAGCTTAAGTTTAAACCGATACCAATGAGACGGAAGGCAT 1026
QY 137 ValPheSerTyrGlyIleIleLeuTyrGluValIleThrArgArgLysProPheAspGlu 156
Db 1027 GTTATAGTATTGGACTTCTTTATGGAAATGTAAGCTGGAGCACTTCCATATAGAGAG 1086
QY 157 IleGlyGlyProAlaPheArgIleMetTyrAlaVal---HisAsnGlyThrArgProPro 175
Db 1087 ATGAATATT---GCTGAACAAATTCCTACCGCATATATACAAGAAATTAAGCCAGTT 1143
QY 176 LeuIleLysAsnLeuProLysProIleGluSerLeuMetThrArgCysTyrSerLysAsp 195
Db 1144 ATACCGACGAGATTGTCAGCGCGCATGAAGAGCTGATGCGATGTTGTGTCATCCGAA 1203
QY 196 ProSerGlnArgProSerMetGluGluIleValLysIleMetThrHisLeuMetArg 214
Db 1204 ACAGACAGAGACCGGAATTCTGGCAGATTGTTCAGAGTGTGGAACTTTTCAAGAG 1260

RESULT 9
US-09-938-842A-1014
; Sequence 1014, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kieps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938, 842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1014
; LENGTH: 1662
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1014

Alignment Scores:
Pred. No.: 2 13e-31 Length: 1662
Score: 345.50 Matches: 80
Percent Similarity: 51.69% Conservative: 42
Best Local Similarity: 33.90% Mismatches: 87
Query Match: 27.60% Indels: 27
DB: 9 Gaps: 8

US-09-830-144-2_COPY_76_303 (1-228) x US-09-938-842A-1014 (1-1662)
QY 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCys 21
|||||
::: ||| |||
|||||
```

```
Db 958 GAAATATATATATAGAGAAAGTTCCGCAATAAAATGTTGTCAGTTCAATGGTCAGTGT 1017
QY 22 LeuAsnPro-----ValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsn 39
Db 1018 ACAGATACCAAAACCTCTGCAATTTGTGACAGAGTTTATGACTCGGGGAGCATTTATGAT 1077
QY 40 ValLeuHisGlyValaGluProLysProTyrTyrThrAlaAlaHisAlaMetSerTyrCys 59
Db 1078 TTCCTTCACAAACCAAGGGGTT-----TTTAAATTCAACTTTTGCTCAAGTGC 1131
QY 60 LeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProlysAlaLeuIleHis 79
Db 1132 CTGAGCTCTCGAAAGAAATGAATTAATTCGATCAA-----AACATATATTCAAT 1182
QY 80 ArgAspLeuLysProProHisLeuLeuValAlaGlyGlyThrValLeuLysIleCys 99
Db 1183 AGAGACTTAAAGCTGTAATCTTCTTANG---GACGAACATGAAGTTGTCAAAGTTGCC 1239
QY 100 AspPheGlyThrAlaCysAspIleGlnThrHis-----MetThrAsnAsnLysGly 116
Db 1240 GATTTTGGTGTGCTCC---AGAGTCAGACTGATGTCAGGGGTTATGCAACGGAACAGGG 1296
QY 117 SerAlaAlaTyrMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAsp 136
Db 1297 ACATACCGATGATGGCTCCAGAGGTCATTAAGACACAAACCTTATGATCAGAGGCGAGAT 1356
QY 137 ValPheSerTyrGlyIleIleLeuTyrGluValIleThrArgArgLysProPheAspGlu 156
Db 1357 GTCTTCAGCTAGCAGATTTGCTGCTGGGAACCTTTGACCTGGGGAACCTCCATATCTTAC 1416
QY 157 IleGlyGlyProAlaPheArgIleMetTyrAlaValHisAsnGlyThrArgProProLeu 176
Db 1417 TTG---ACTCCATGCGAACCTGCTGTGGCTGTGCCAAAGGACTTAAGACCAAAATT 1473
QY 177 IleLysAsnLeuProLysProIleGluSerLeuMetThrArgCysTyrSerLysAspPro 196
Db 1474 CCAAAGGAAACACCCCAAACTGACTGAACTTTTGAGAAATGGCTGGAGCAAGACCA 1533
QY 197 SerGlnArgProSerMetGluGluIleValLysIleMetThrHisLeuMetArg----- 214
Db 1534 GCTCTAAGACCAAAATTTGCAGAAATCATAGAAATGCTTAACCAACTAATCGCAGAGTA 1593
QY 215 -----TyrPheProGly 218
Db 1594 ATGATATTATCATTCGATTAAGATTAACATGATGGTGTCTTTCAGCG 1641

RESULT 10
US-09-947-199-3
; Sequence 3, Application US/09947199
; Patent No. US20020127684A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jayaseelan
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MNI-068CP2
; CURRENT APPLICATION NUMBER: US/09/947,199
; PRIOR FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: 60/111,938
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/291,839
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 09/458,457
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2505
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: CDS
; LOCATION: (1)..(2505)
US-09-947-199-3
```

```
Alignment Scores:
Pred. No.: 2,61e-30 Length: 2505
Score: 338.50 Matches: 84
Percent Similarity: 59.07% Conservative: 43
Best Local Similarity: 39.07% Mismatches: 72
Query Match: 27.04% Indels: 17
DB: 10 Gaps: 8

US-09-830-144-2_COPY_76_303 (1-228) x US-09-947-199-3 (1-2505)

QY 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCys 21
Db 1525 GAGGTGTCATCTCTGCCAGCTCAATCATCCCTGCGTAATTCAGTTGTGGGTGCTTGC 1584
QY 22 LeuAsn-----ProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyr 38
Db 1585 TTGAATGATCCAGCCAGTTGCCATTGTCTCACTCAATACATATCAGGGGGTCTCTGTTTC 1644
QY 39 AsnValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrp 58
Db 1645 TCCCTCCTTCATGA-GCAGAAGAGGATTCTTGATTT-----GCAGTCTAAATAATTAT 1697
QY 59 CysLeuGlnCysSerGln-GlyValAlaTyrLeuHisSerMet---GlnProLysAlaLe 77
Db 1698 TGCAGTAGATGTTGCCAAGGCGATGGAGTACCTTCAACACTGACACAGCCA-----AT 1751
QY 77 uIleHisArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLy 97
Db 1752 TATACATCGTGACTTGAACAGTCAATATCTTCTCTATGAGGTGGGCTGCTGTG-- 1809
QY 97 sIleCysAspPheGlyThrAlaCysAspIleGlnThr-----HisMetThrAs 113
Db 1810 -GTGGCAGATTTTGCAGAAATCAAGATTCTACAGTCTCTGAGTGAAGACAAATGACAAA 1868
QY 113 nAsnLysGlySerAlaAlaTrpMetAlaProGluValPhe---GluGlySerAsnTyrSe 132
Db 1869 ACAACCTGGGAACCTCGTTGGATGGCTCTGAGGTGTTTCCAGCAGTGCACTCGGTACAC 1928
QY 132 rGluLysCysAspValPheSerTrpGlyIleIleLeuTrpGluValIleThrArgArgLy 152
Db 1929 CATCAAGCAGATGTTCTCAGCTATGCTGTGTCTGTGGAAATCTCTACTGGCGAAAT 1988
QY 152 sProPheAspGluIleGlyGlyProAlaPheArgIleMetTrpAlaValHisAsnGlyTh 172
Db 1989 TCCATTGCTCATCTCAAGCCAGCGGCTGCGGAGCAGACATGCTTACCACCAC---AT 2045
QY 172 rArgProProLeuLysAsnLeuProLysProIleGluSerLeuMetThrArgCysTr 192
Db 2046 CAGACCTCCCATTTGGCTATTTCATTCCTCAAGCCCATATCATCTCTGCTGATACGAGGTG 2105
QY 192 pSerLysAspProSerGlnArgProSerMetGluGluIleVal 206
Db 2106 GAACGCATGTCCTGAAGAGACCCGAATTTTCTGAAGTTGTC 2148

RESULT 11
US-09-947-199-1
; Sequence 1, Application US/09947199
; Patent No. US20020127684A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: THEREFOR
; CURRENT APPLICATION NUMBER: US/09/947,199
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: 60/111,938
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/291,839
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 09/458,457
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 9
```

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3025
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (48)..(2552)
US-09-947-199-1

Alignment Scores:
Pred. No.: 3,43e-30 Length: 3025
Score: 338.50 Matches: 84
Percent Similarity: 59.07% Conservative: 43
Best Local Similarity: 39.07% Mismatches: 72
Query Match: 27.04% Indels: 17
DB: 10 Gaps: 8

US-09-830-144-2_COPY_76_303 (1-228) x US-09-947-199-1 (1-3025)

QY 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCys 21
Db 1572 GAGGTGTCATCTCTGCCAGCTCAATCATCCCTGCGTAATTCAGTTGTGGGTGCTTGC 1631
QY 22 LeuAsn-----ProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyr 38
Db 1632 TTGAATGATCCAGCCAGTTGCCATTGTCTCACTCAATACATATCAGGGGGTCTCTGTTTC 1691
QY 39 AsnValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrp 58
Db 1692 TCCCTCCTTCATGA-GCAGAAGAGGATTCTTGATTT-----GCAGTCTAAATAATTAT 1744
QY 59 CysLeuGlnCysSerGln-GlyValAlaTyrLeuHisSerMet---GlnProLysAlaLe 77
Db 1745 TGCAGTAGATGTTGCCAAGGCGATGGAGTACCTTCAACACTGACACAGCCA-----AT 1798
QY 77 uIleHisArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLy 97
Db 1799 TATACATCGTGACTTGAACAGTCAATATTTCTTCTCTATGAGGTGGGCTGCTGTG-- 1856
QY 97 sIleCysAspPheGlyThrAlaCysAspIleGlnThr-----HisMetThrAs 113
Db 1857 -GTGGCAGATTTTGCAGAAATCAAGATTCTTACAGTCTCTGAGTGAAGACAAACATGACAAA 1915
QY 113 nAsnLysGlySerAlaAlaTrpMetAlaProGluValPhe---GluGlySerAsnTyrSe 132
Db 1916 ACAACCTGGGAACCTCGTTGGATGGCTCTGAGGTGTTTCCAGCAGTGCACTCGGTACAC 1975
QY 132 rGluLysCysAspValPheSerTrpGlyIleIleLeuTrpGluValIleThrArgArgLy 152
Db 1976 CATCAAGCAGATGTTCTCAGCTATGCTGTGTCTGTGGGAAATTTCTCACTGGCGAAAT 2035
QY 152 sProPheAspGluIleGlyGlyProAlaPheArgIleMetTrpAlaValHisAsnGlyTh 172
Db 2036 TCCATTGCTCATCTCAAGCCAGCGGCTGCGGAGCAGACATGCTTACCACCAC---AT 2092
QY 172 rArgProProLeuLysAsnLeuProLysProIleGluSerLeuMetThrArgCysTr 192
Db 2093 CAGACCTCCCATTTGGCTATTTCATTCCTCAAGCCCATATCATCTCTGCTGATACGAGGTG 2152
QY 192 pSerLysAspProSerGlnArgProSerMetGluGluIleVal 206
Db 2153 GAACGCATGTCCTGAAGAGACCCGAATTTTCTGAAGTTGTC 2195

RESULT 12
US-09-947-199-9
; Sequence 9, Application US/09947199
; Patent No. US20020127684A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: THEREFOR
; FILE REFERENCE: MNI-068CP2
```

```

; CURRENT APPLICATION NUMBER: US/09/947,199
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: 60/111,938
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/291,839
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 09/458,457
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 2505
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2505)
US-09-947-199-9
```

```

Alignment Scores:
Pred. No.: 5,9e-30 Length: 2505
Score: 335.50 Matches: 80
Percent Similarity: 56.54% Conservative: 41
Best Local Similarity: 37.38% Mismatches: 78
Query Match: 26.80% Indels: 15
DB: 10 Gaps: 8
```

US-09-830-144-2_COPY_76_303 (1-228) x US-09-947-199-9 (1-2505)

```

QY 2 GtLueuArgGlnLeuSerArgValaAenHisProAsnIleValysLeuTyrgIyalaCys 21
Db 1525 GAGGTGTCAATTCCTGCGACGTCACACCCCTGCGTGGTTCAGTTGGTGCCCTGC 1584
QY 22 Leu-----AsnProValCysLeuValMetGluTyrgIyalaGluGlySerLeuTy 38
Db 1585 CTGGATGACCCCGACGATTTGGCATTCCTGCTCACTCACTGATTCAGAGGCTCCCTGTT 1644
QY 39 AsnValLeuHisGlyValaGluProLeuProTyTyThraIalaHisIaMetSerTrp 58
Db 1645 TCCTGCTTCAT-----GAACAGAAAGAAATTCCTGACTTCAGTCTAAATTAATTCAT 1698
QY 59 CysLeuGlnCysSerGlnGlyValaIaTyrlLeuHisSerMet---GlnProlysaLeu 77
Db 1699 GCGGTAGACGTTGCCAAGGCGCATGAGTACCTGCACAGCTTACCCAGCA-----ATC 1752
QY 78 IleHisArgAspLeuLysProProAsnLeuLeuValaIaGlyTyThraValLeuLys 97
Db 1753 ATACACCGGACCTGGAACGCCACAAATATTCCTGCTATAGAGATGGCCATGCTGTG--- 1809
QY 98 IleCysAspPheGlyThraIaCysAspIleGlnThr-----HisMetThrAsn 113
Db 1810 GTGGCAGATTTGGAGAAATCAAGATTCGTCAGTCCCTGATGAAGACAAATGACAAAG 1869
QY 114 AsnLysGlySerAlaIaIaTrpMetAlaProGluValaPhe---GluGlySerAsnTy 132
Db 1870 CAGCCAGGAACCTGCGCGTGGATGGCCCTGAGGTGTACACAGGCGACGAGATACCC 1929
QY 133 GluLysCysAspValaPheSerTrpGlyIleIleLeuTyrgIyalaIleThraArgLys 152
Db 1930 ATCAAGCGTGATCTTCACTGATTAATCCCTGCTGTGGAGCTCCTCACTGAGAAAT 1989
QY 153 ProPheAspGluIleGlyLysProAlaPheArgIleMetTrpAlaValHisAsnGlyThr 172
Db 1990 CCATTGCTCATCTCAAGCAGCGCTGCAGACAGATATGCGGATACACCA---ATC 2046
QY 173 ArgProProLeuIleLysAsnLeuProlysaProIleGluSerLeuMetThrArgCysTrp 192
Db 2047 AGACCCCGCATGCGCTAATTCATCCCAAGCCCATCTCATCCCTGCTGTAGAGGGGCTGG 2106
QY 193 SerLysAspProSerGlnArgProSerMetGluGluIleVal 206
Db 2107 AATGATGTCTCTGAAGAGCAGCCAGAGTTCTCTGAAGTGCTT 2148
```

```

RESULT 13
US-09-947-199-7
; Sequence 7, Application US/09947199
; Patent No. US20020127684A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jayaseelan
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE OF INVENTION: THEREFOR
; FILE REFERENCE: MNI-068CP2
; CURRENT APPLICATION NUMBER: US/09/947,199
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: 60/111,938
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/291,839
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 09/458,457
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 3026
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)..(2565)
US-09-947-199-7
```

```

Alignment Scores:
Pred. No.: 7,77e-30 Length: 3026
Score: 335.50 Matches: 80
Percent Similarity: 56.54% Conservative: 41
Best Local Similarity: 37.38% Mismatches: 78
Query Match: 26.80% Indels: 15
DB: 10 Gaps: 8
```

US-09-830-144-2_COPY_76_303 (1-228) x US-09-947-199-7 (1-3026)

```

QY 2 GtLueuArgGlnLeuSerArgValaAenHisProAsnIleValysLeuTyrgIyalaCys 21
Db 1585 GAGGTGTCAATTCCTGCGACGTCACACCCCTGCGTGGTTCAGTTGGTGCCCTGC 1644
QY 22 Leu-----AsnProValCysLeuValMetGluTyrgIyalaGluGlySerLeuTy 38
Db 1645 CTGGATGACCCCGACGATTTGGCATTCCTGCTCACTCACTGATTCAGAGGCTCCCTGTT 1704
QY 39 AsnValLeuHisGlyValaGluProLeuProTyTyThraIalaHisIaMetSerTrp 58
Db 1705 TCCTGCTTCAT-----GAACAGAAAGAAATTCCTGACTTCGACGTCATAATTAATCAT 1758
QY 59 CysLeuGlnCysSerGlnGlyValaIaTyrlLeuHisSerMet---GlnProlysaLeu 77
Db 1759 GCGGTAGACGTTGCCAAGGCGCATGAGTACCTGCACAGCTTACCCAGCA-----ATC 1812
QY 78 IleHisArgAspLeuLysProProAsnLeuLeuValaIaGlyTyThraValLeuLys 97
Db 1813 ATACACCGGACCTGGAACGCCACAAATATTCCTGCTATAGAGATGGCCATGCTGTG--- 1869
QY 98 IleCysAspPheGlyThraIaCysAspIleGlnThr-----HisMetThrAsn 113
Db 1870 GTGGCAGATTTGGAGAAATCAAGATTCGTCAGTCCCTGATGAAGACAAATGACAAAG 1929
QY 114 AsnLysGlySerAlaIaIaTrpMetAlaProGluValaPhe---GluGlySerAsnTy 132
Db 1930 CAGCCAGGAACCTGCGCGTGGATGGCCCTGAGGTGTACACAGGCGACGAGATACACC 1989
QY 133 GluLysCysAspValaPheSerTrpGlyIleIleLeuTyrgIyalaIleThraArgLys 152
Db 1990 ATCAAGCGTGATCTTCACTGATTAATCCCTGCTGTGGAGCTCCTCACTGAGAAAT 2049
QY 153 ProPheAspGluIleGlyLysProAlaPheArgIleMetTrpAlaValHisAsnGlyThr 172
Db 2050 CCATTGCTCATCTCAAGCAGCGCTGCAGACAGATATGCGGATACACCA---ATC 2106
```


DR WPI: 1997-515318/48.
 DR N-PSDB; AAT91175.
 PT DNA encoding TAK1 binding protein TAB1 - member of transforming
 PT growth factor beta receptor signal production pathway, which
 PT activates TAK-1 kinase activity upon binding
 XX
 XX
 PS Example 5; Page 17-19; 30pp; English.
 CC This protein comprises human TAB1, a novel member of the
 CC transforming growth factor-beta receptor signal transduction
 CC pathway, which activates TAK-1 kinase activity upon binding. Its
 CC amino acid sequence was deduced from a cDNA clone (see AAT91175)
 CC obtained from a kidney library; a variant TAB1 (see AAW2707) has
 CC Arg rather than Ser at amino acid position 52. Also claimed are:
 CC isolated DNA encoding a protein modified by a substitution,
 CC deletion and/or addition of 1 or more amino acids of the 504-residue
 CC TAB1 sequence; (2) DNA which can hybridise with the 1560 bp TAB1
 CC nucleic acid sequence; (3) isolated DNA encoding a protein
 CC comprising amino acids 21-579 or 437-504 of the 504 TAB1 sequence;
 CC (4) DNA encoding a fusion protein comprising an above protein or
 CC polypeptide; (5) expression vector comprising an above DNA; and
 CC (6) host cell, preferably a mammalian or yeast cell, transformed by
 CC the expression vector. Cells expressing TAB1 and TAK1 can be used
 CC to screen for TGF-beta signalling pathway inhibitors by contacting
 CC the cells with a test compound, and measuring the TAK1 kinase
 CC activity.
 XX
 XX
 SQ Sequence 504 AA:
 Query Match 100.0%; Score 2580; DB 18; Length 504;
 Best Local Similarity 100.0%; Pred. No. 1e-212;
 Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ID AAY09541 standard; Protein; 504 AA.
 XX
 AC AAY09541.
 XX
 XX
 DT 21-JUL-1999 (first entry)
 XX
 XX
 DE Human TAB1 protein.
 XX
 XX
 KW Human; TAB1; screening; inhibition; TGF-beta;
 KW transforming growth factor beta.
 XX
 XX
 OS Homo sapiens.
 XX
 PN W09921010-A1.
 XX
 PD 29-APR-1999.
 XX
 PF 22-OCT-1998; 98WO-JP04796.
 XX
 PR 22-OCT-1997; 97JP-0290188.
 XX
 PA (CHUS) CHUGAI SEIYAKU KK.
 XX
 PI Ohtomo T, Ono K, Tsuchiya M;
 XX
 DR WPI: 1999-312645/26.
 DR N-PSDB; AAX56278.
 XX
 PT Screening for TGF-beta inhibitory substances, which are useful as
 PT drugs for treatment of diseases relating to its disorder
 XX
 PS Claim 3; Page 147-149; 195pp; Japanese.
 XX
 XX
 CC A method has been developed for screening for substances which inhibit
 CC the binding of TAK1 polypeptide to TAB1 polypeptide. The method
 CC comprises: (a) contacting the polypeptide in the presence of a sample;
 CC and (b) detecting the amount of bound polypeptide, in which the sample
 CC can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming
 CC growth factor (TGF)-beta inhibitory substances can be used in drugs for
 CC indications e.g. as TGF-beta signal transduction inhibitors or
 CC activators, or extracellular matrix protein production enhancement
 CC inhibitors, or cell proliferation prevention inhibitors or
 CC activators, or monocytic migration inhibitors or activators, or
 CC physiological activity induction inhibitors or activators, or
 CC immunosuppression inhibitors or activators, or amyloid beta protein
 CC precipitation inhibitors or activators, and such substances can also be
 CC inhibitors of the TAK1 polypeptide function, particularly kinase
 CC activity. The present sequence represents human TAB1.
 XX
 XX
 SQ Sequence 504 AA:
 Query Match 100.0%; Score 2580; DB 20; Length 504;
 Best Local Similarity 100.0%; Pred. No. 1e-212;
 Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 241 RRIGDYKVKYGYTDIDLLSAAKSPKIIAEPEIHGAQPLDGVGTGFLVLMSEGLYKALEAAH 300
Qy 301 GPGQANOEIAAMIDTEFAKOTSLDAVAQVVDVRVKRIHSDTFASGGERARFCPRHEDMTL 360
Db 301 GPGQANOEIAAMIDTEFAKOTSLDAVAQVVDVRVKRIHSDTFASGGERARFCPRHEDMTL 360
Qy 361 LVRNFGYPLGEMSOPTSPAPAGRVYPVSPYSSAQSTKTSVTLSLVMPSQOGQWVNG 420
Db 361 LVRNFGYPLGEMSOPTSPAPAGRVYPVSPYSSAQSTKTSVTLSLVMPSQOGQWVNG 420
Qy 421 AHSASTLDEATPTLTNOSPRTLLOSTNTHQTQSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480
Db 421 AHSASTLDEATPTLTNOSPRTLLOSTNTHQTQSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480
Qy 481 YVDFAEFYRLMSVDHGEQSVVTAP 504
Db 481 YVDFAEFYRLMSVDHGEQSVVTAP 504
RESULT 3
AAV91001
ID AAV91001 standard; Protein; 504 AA.
XX
AC AAV91001;
XX
DT 04-SEP-2000 (first entry)
XX Human TAB-1 protein sequence SEQ ID NO:4.
XX
XX Human; TAK-1; TAB-1; mitogen activated protein kinase; MAPK;
KW screening; signal transduction; inhibition; inflammatory cytokine;
KW IL-1; interleukin 1; TNF; tumour necrosis factor; inflammation;
XX antiinflammatory; suppression.
OS Homo sapiens.
XX
XX W0200023610-A1.
XX
XX 27-APR-2000.
PD
XX
XX 21-OCT-1999; 99WO-JP05817.
PF
XX
XX 21-OCT-1998; 98JP-0299962.
PR
XX
XX (CHUS) CHUGAI SEIYAKU KK.
XX
XX Tsuchiya M, Ontomo T, Sugamata Y, Matsumoto K;
PI
XX WPI; 2000-339707/29.
XX
XX N-PSDB; AAA39106.
DR
XX
XX Method for screening inhibitors of TAK1 signal transduction for
PT suppression of inflammatory cytokine production and use as
PT antiinflammatory agents -
XX
XX Disclosure; Page 90-94; 100pp; Japanese.
XX
XX The present invention describes a method for screening compounds for
CC inhibition of inflammatory cytokine signal transduction by contacting
CC the sample with TAK1 and its receptor TAB1 and selecting for inhibition
CC of TAK1/TAB1 binding. Also described is a method for screening compounds
CC for inhibition of inflammatory cytokine signal transduction in which the
CC inhibition of TAK1 phosphorylation is selected for; and drug
CC compositions for the treatment of inflammatory disorders containing as
CC active component an inflammatory cytokine signal transduction inhibitor.
CC TAK1 is an essential component of the signalling process which results
CC in release of inflammatory cytokines such as interleukin-1 (IL-1),
CC IL-10, tumour necrosis factor (TNF) and IL-6. The methods can be used
CC for the selection of effective antiinflammatory agents. The present
CC sequence represents human TAB-1, which is used in the exemplification of
CC the present invention.
XX
XX Sequence 504 AA;
SQ

Query Match 100.0%; Score 2580; DB 21; Length 504;
Best Local Similarity 100.0%; Pred. No. le-212;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAQRRLSLQSEQPSWTDLDPLCHLSGVGSASNRYSADGKGTESHPEDSWLKFRSEN 60
Db 1 MAQRRLSLQSEQPSWTDLDPLCHLSGVGSASNRYSADGKGTESHPEDSWLKFRSEN 60
Qy 61 NCFLYGVNFGYDGNVTNFAQRLSAELLGQLNAEHAEDVRRVLLQAFDVERSFLS 120
Db 61 NCFLYGVNFGYDGNVTNFAQRLSAELLGQLNAEHAEDVRRVLLQAFDVERSFLS 120
Qy 121 IDDAEAERKASLOSQPEGVPOHLPPOYKILERLKTLEIREISGGAMAVVLLNNKLYV 180
Db 121 IDDAEAERKASLOSQPEGVPOHLPPOYKILERLKTLEIREISGGAMAVVLLNNKLYV 180
Qy 181 ANVGTNRALLCKSTVDGLQVTQLNVDTHTTENEDELFRLSQLGLDAGKIKQVGIIICGQEST 240
Db 181 ANVGTNRALLCKSTVDGLQVTQLNVDTHTTENEDELFRLSQLGLDAGKIKQVGIIICGQEST 240
Qy 241 RRIGDYKVKYGYTDIDLLSAAKSKPIIAEPRIHGAQPLDGVGTGFLVLMSEGLYKALEAAH 300
Db 241 RRIGDYKVKYGYTDIDLLSAAKSKPIIAEPRIHGAQPLDGVGTGFLVLMSEGLYKALEAAH 300
Qy 301 GPGQANOEIAAMIDTEFAKOTSLDAVAQVVDVRVKRIHSDTFASGGERARFCPRHEDMTL 360
Db 301 GPGQANOEIAAMIDTEFAKOTSLDAVAQVVDVRVKRIHSDTFASGGERARFCPRHEDMTL 360
Qy 361 LVRNFGYPLGEMSOPTSPAPAGRVYPVSPYSSAQSTKTSVTLSLVMPSQOGQWVNG 420
Db 361 LVRNFGYPLGEMSOPTSPAPAGRVYPVSPYSSAQSTKTSVTLSLVMPSQOGQWVNG 420
Qy 421 AHSASTLDEATPTLTNOSPRTLLOSTNTHQTQSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480
Db 421 AHSASTLDEATPTLTNOSPRTLLOSTNTHQTQSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480
Qy 481 YVDFAEFYRLMSVDHGEQSVVTAP 504
Db 481 YVDFAEFYRLMSVDHGEQSVVTAP 504
RESULT 4
AAV59450
ID AAV59450 standard; Protein; 504 AA.
XX
AC AAV59450;
XX
DT 24-MAR-2000 (first entry)
XX
XX Human TAB1 protein sequence.
XX
XX Human; TAB1; XIAP; X-linked inhibitor of apoptosis protein; TGF-beta;
KW transforming growth factor-beta activated kinase 1; monocyte migration;
KW TAK1 binding protein 1; extracellular matrix protein production;
KW cell growth inhibitor; beta-amyloid protein deposition;
KW immunosuppression; Transforming growth factor-beta.
XX
XX Homo sapiens.
XX
XX JP11326328-A.
PN
XX
XX 26-NOV-1999.
PD
XX
XX 13-MAY-1998; 98JP-0130378.
PF
XX
XX 13-MAY-1998; 98JP-0130378.
PR
XX
XX (MATS/) MATSUMOTO K.
PA
XX
XX WPI; 2000-078337/07.
DR
XX
XX N-PSDB; AAZ48861.
DR
XX

PT Screening a substance which inhibits combination of the X-linked
PT inhibitor of apoptosis protein -
XX
XX
PS Claim 2; Page 25-26; 43pp; Japanese.
XX
CC This sequence represents the human TAB1 protein.
CC The invention relates to a method for screening a substance inhibiting
CC the formation of a complex between XIAP and TAB1, in which X-linked
CC inhibitor of apoptosis protein (XIAP), transforming growth factor-beta
CC activated kinase 1 (TAK1) binding protein 1 (TAB1) and a substance to be
CC tested are contacted with each other and then the presence or formation
CC of a complex between XIAP and TAB1 is detected. The substance can be used
CC as a drug for extracellular matrix protein production enhancement, cell
CC growth inhibition, monocyte migration, physiologically active substance
CC induction, immunosuppression, and beta-amyloid protein deposition. A
CC substance inhibiting the formation of a complex between TAB1 and XIAP as
CC well as between XIAP and TGF-beta (Transforming growth factor-beta) type
CC I and/or type II receptor is useful as a drug.
XX
SQ Sequence 504 AA;
Query Match 100.0%; Score 2580; DB 21; Length 504;
Best Local Similarity 100.0%; Pred. No. 1e-212;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAQRSLQSEQPSWTDLPLCHLSGVGSASNSRSYADGKGTESHPPEDSWLKFRSEN 60
DB 1 MAQRSLQSEQPSWTDLPLCHLSGVGSASNSRSYADGKGTESHPPEDSWLKFRSEN 60
QY 61 NCFLYGVFNQYDGNRVNTPFAQRLSAELLIGQLNAEHAADVRRVLLQAFDVVERSFLES 120
DB 61 NCFLYGVFNQYDGNRVNTPFAQRLSAELLIGQLNAEHAADVRRVLLQAFDVVERSFLES 120
QY 121 IDALAERKASLQSLPEGVPOHLPPOYOKILERLKLTEREISGGAMAAVAVLLNNKLYV 180
DB 121 IDALAERKASLQSLPEGVPOHLPPOYOKILERLKLTEREISGGAMAAVAVLLNNKLYV 180
QY 181 ANVTNRALLCKSTVDGLQVLTQNVNHTTENEDLFRLSQLGIDAGKIKOVGIIICQEST 240
DB 181 ANVTNRALLCKSTVDGLQVLTQNVNHTTENEDLFRLSQLGIDAGKIKOVGIIICQEST 240
QY 241 RRIQDYKVKYGYTDIDLLSAKSKPIIAEPEIHGAQPLDVGTFVLVMSGLYKALEAAH 300
DB 241 RRIQDYKVKYGYTDIDLLSAKSKPIIAEPEIHGAQPLDVGTFVLVMSGLYKALEAAH 300
QY 301 GPQANQEIAMIDTEFAKQTSIDAVAQAVDVKRIHSDTFASGGERARFCPRHEDMTL 360
DB 301 GPQANQEIAMIDTEFAKQTSIDAVAQAVDVKRIHSDTFASGGERARFCPRHEDMTL 360
QY 361 LVNFGYPLGEMSQPTPSAPAAAGRVYPVSSAQSTSKTSVTLSLMP9QGOMVNG 420
DB 361 LVNFGYPLGEMSQPTPSAPAAAGRVYPVSSAQSTSKTSVTLSLMP9QGOMVNG 420
QY 421 AHSASTLDEATPPLTNOSPTLLOSTNTHOSSSSDGLFPRSPAHSLPGEEDGRVAP 480
DB 421 AHSASTLDEATPPLTNOSPTLLOSTNTHOSSSSDGLFPRSPAHSLPGEEDGRVAP 480
QY 481 YVDFAEFYRLMSVDHGQSVVTAP 504
DB 481 YVDFAEFYRLMSVDHGQSVVTAP 504
RESULT 5
AA09550
ID AA09550 standard; protein; 513 AA.
XX
AC AA09550;
XX
DT 21-JUL-1999 (first entry)
XX
DE Human TAB1 protein SEQ ID NO:43.
XX
KW Human; TAB1; TAK1; screening; inhibition; TGF-beta;

KW transforming growth factor beta.
XX
XX Homo sapiens.
OS
XX
PN WO9921010-A1.
XX
PD 29-APR-1999.
XX
PF 22-OCT-1998; 98MO-JP04796.
XX
PR 22-OCT-1997; 97JP-0290188.
XX
PA (CHUS) CHUGAI SEIYAKU KK.
XX
PI Ohtomo T, Ono K, Tsuchiya M;
XX
DR WPI; 1999-312645/26.
XX
PT N-PSDB; AAX56310.
XX
PT Screening for TGF- beta inhibitory substances, which are useful as
PT drugs for treatment of diseases relating to its disorder
XX
PS Example 13; Page 186-188; 195pp; Japanese.
XX
CC A method has been developed for screening for substances which inhibit
CC the binding of TAK1 polypeptide to TAB1 polypeptide. The method
CC comprises: (a) contacting the polypeptide in the presence of a sample;
CC and (b) detecting the amount of bound polypeptide, in which the sample
CC can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming
CC growth factor (TGF)-beta inhibitory substances can be used in drugs for
CC indications e.g. as TGF-beta signal transduction inhibitors or
CC activators, or extracellular matrix protein production enhancement
CC inhibitors, or activators, or cell proliferation prevention inhibitors or
CC activators, or monocyte migration inhibitors or activators, or
CC physiological activity induction inhibitors or activators, or
CC immunosuppression inhibitors or activators, or amyloid beta protein
CC precipitation inhibitors or activators, and such substances can also be
CC inhibitors of the TAK1 polypeptide function, particularly kinase
CC activity. The present sequence represents human TAB1.
XX
SQ Sequence 513 AA;
Query Match 100.0%; Score 2580; DB 20; Length 513;
Best Local Similarity 100.0%; Pred. No. 1e-212;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAQRSLQSEQPSWTDLPLCHLSGVGSASNSRSYADGKGTESHPPEDSWLKFRSEN 60
DB 1 MAQRSLQSEQPSWTDLPLCHLSGVGSASNSRSYADGKGTESHPPEDSWLKFRSEN 69
QY 61 NCFLYGVFNQYDGNRVNTPFAQRLSAELLIGQLNAEHAADVRRVLLQAFDVVERSFLES 120
DB 70 NCFLYGVFNQYDGNRVNTPFAQRLSAELLIGQLNAEHAADVRRVLLQAFDVVERSFLES 129
QY 121 IDALAERKASLQSLPEGVPOHLPPOYOKILERLKLTEREISGGAMAAVAVLLNNKLYV 180
DB 130 IDALAERKASLQSLPEGVPOHLPPOYOKILERLKLTEREISGGAMAAVAVLLNNKLYV 189
QY 181 ANVTNRALLCKSTVDGLQVLTQNVNHTTENEDLFRLSQLGIDAGKIKOVGIIICQEST 240
DB 190 ANVTNRALLCKSTVDGLQVLTQNVNHTTENEDLFRLSQLGIDAGKIKOVGIIICQEST 249
QY 241 RRIQDYKVKYGYTDIDLLSAKSKPIIAEPEIHGAQPLDVGTFVLVMSGLYKALEAAH 300
DB 250 RRIQDYKVKYGYTDIDLLSAKSKPIIAEPEIHGAQPLDVGTFVLVMSGLYKALEAAH 309
QY 301 GPQANQEIAMIDTEFAKQTSIDAVAQAVDVKRIHSDTFASGGERARFCPRHEDMTL 360
DB 310 GPQANQEIAMIDTEFAKQTSIDAVAQAVDVKRIHSDTFASGGERARFCPRHEDMTL 369
QY 361 LVNFGYPLGEMSQPTPSAPAAAGRVYPVSSAQSTSKTSVTLSLMP9QGOMVNG 420
DB 370 LVNFGYPLGEMSQPTPSAPAAAGRVYPVSSAQSTSKTSVTLSLMP9QGOMVNG 429

PS Example 5; Page 19-21; 30pp; English.
XX This protein comprises human TAB1, a novel member of the
CC transforming growth factor-beta receptor signal transduction
CC pathway, which activates TAK-1 kinase activity upon binding. Its
CC amino acid sequence was deduced from a cDNA clone (see AAT91176)
CC obtained from a kidney library; a variant TAB1 (see AAW26706) has
CC ser rather than Arg at amino acid position 52. Also claimed are:
CC isolated DNA encoding a protein modified by a substitution,
CC deletion and/or addition of 1 or more amino acids of the 504-residue
CC TAB1 sequence; (2) DNA which can hybridise with the 1560 bp TAB1
CC nucleic acid sequence; (3) isolated DNA encoding a protein
CC comprising amino acids 21-579 or 437-504 of the 504 TAB1 sequence;
CC (4) DNA encoding a fusion protein comprising an above protein or
CC polypeptide; (5) expression vector comprising an above DNA; and
CC (6) host cell, preferably a mammalian or yeast cell, transformed by
CC the expression vector. Cells expressing TAB1 and TAK1 can be used
CC to screen for TGF-beta signalling pathway inhibitors by contacting
CC the cells with a test compound, and measuring the TAK1 kinase
CC activity.
CC
CC
CC
CC
SQ Sequence 504 AA;

Query Match 99.8%; Score 2575; DB 18; Length 504;
Best Local Similarity 99.8%; Pred. No. 2,7e-212;
Matches 503; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MAQRASLQSEQSPWTDLPLCHLSGVGSASNRYSADGKTESHPEDSMLKFRSEN 60
DB 1 MAQRASLQSEQSPWTDLPLCHLSGVGSASNRYSADGKTESHPEDSMLKFRSEN 60
OY 61 NCFLYGVNGYDGNRVTFNVAORLSAELLIGQLNMEHAEADVRRVLQAFDVERSFLBS 120
DB 61 NCFLYGVNGYDGNRVTFNVAORLSAELLIGQLNMEHAEADVRRVLQAFDVERSFLBS 120
OY 121 IDALAEKASLSQSPREGVPOHQLPROYOKILRLKTEREISGGMAVAVALLNNKLYV 180
DB 121 IDALAEKASLSQSPREGVPOHQLPROYOKILRLKTEREISGGMAVAVALLNNKLYV 180
OY 181 ANVTNRALLCKSTVDGLQVTLQNVHTTENEDLFLRLSQLGDAGKIKQVGIIQGEEST 240
DB 181 ANVTNRALLCKSTVDGLQVTLQNVHTTENEDLFLRLSQLGDAGKIKQVGIIQGEEST 240
OY 241 RRIQGVKXVGYTIDLLLSAASKPIIAPEPIHGAQPLDGVTFVLWSEGLYKALEAH 300
DB 241 RRIQGVKXVGYTIDLLLSAASKPIIAPEPIHGAQPLDGVTFVLWSEGLYKALEAH 300
OY 301 GPGQANQETIAAMIDTEFAKQTSIDAVAQAVVDRVKRIHSDTFASGGERARPCRHEDMTL 360
DB 301 GPGQANQETIAAMIDTEFAKQTSIDAVAQAVVDRVKRIHSDTFASGGERARPCRHEDMTL 360
OY 361 LVNFGPIGEMSQTPSPAPAGRVYPSVYSAQSTKTSVTLSLVMSQCGQWNG 420
DB 361 LVNFGPIGEMSQTPSPAPAGRVYPSVYSAQSTKTSVTLSLVMSQCGQWNG 420
OY 421 AHSASTLDEATPTLTNQSPTLTLTQSTNTHTQSSSSSDGGLFRSRPAHSLPGEDEGRVDP 480
DB 421 AHSASTLDEATPTLTNQSPTLTLTQSTNTHTQSSSSSDGGLFRSRPAHSLPGEDEGRVDP 480
OY 481 YVDFAEFYRLMSVDHGQSIVTAP 504
DB 481 YVDFAEFYRLMSVDHGQSIVTAP 504

RESULT 8
AAB56692
ID AAB56692 standard; Protein; 84 AA.
XX
AC AAB56692;
XX
DT 13-MAR-2001 (first entry)
XX
DE Human prostate cancer antigen protein sequence SEQ ID NO:1270.

XX Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KW neuroprotective; cytosolic; cardioactive; immunomodulatory; muscular;
KW vlninray; gastrointestinal; nephrotoxic; antiinfective; gynaecological;
KW antibacterial; gene therapy; neural; immune; reproductive; renal;
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
KW wound; infectious disease.
XX
OS Homo sapiens.
PN WO200055174-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000NO-US05988.
XX
PR 12-MAR-1999; 99US-0124270.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2000-587513/55.
XX
DR N-PSDB; AAF15895.
XX
PT Prostate cancer associated gene sequences, referred to as prostate
PT cancer antigens, useful for treatment, prevention, and diagnosis of
PT disorders such as prostate cancer -
XX
PS Claim 11; Page 1693; 2338pp; English.

XX AAF15566 to AAF16505 encode the human prostate cancer associated
CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC The prostate cancer antigens can have neuroprotective, cytosolic,
CC cardioactive, immunomodulatory, muscular, vlninray, gastrointestinal,
CC nephrotoxic, antiinfective, gynaecological and antibacterial activities,
CC and can be used in gene therapy. The prostate cancer antigen
CC polynucleotides may be used for detection of prostate cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The prostate cancer antigens may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
CC AAB57303 represent sequences used in the exemplification of the present
CC invention.
CC
CC
CC
SQ Sequence 84 AA;

Query Match 15.9%; Score 409; DB 21; Length 84;
Best Local Similarity 97.5%; Pred. No. 1,9e-27;
Matches 78; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 425 STLDEATPTLTNQSPTLTLTQSTNTHTQSSSSSDGGLFRSRPAHSLPGEDEGRVDP 484
DB 5 ATLDEATPTLTNQSPTLTLTQSTNTHTQSSSSSDGGLFRSRPAHSLPGEDEGRVDP 484
OY 485 AEFYRLMSVDHGQSIVTAP 504
DB 65 AEFYRLMSVDHGQSIVTAP 84

RESULT 9
AAB28132
ID AAB28132 standard; Peptide; 70 AA.
XX
AC AAB28132;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human peptide #783 encoded by breast cell single exon nucleic acid probe.
XX
KW Human; microarray; single exon probe; gene expression; breast;


```

KW disease; cancer.
XX
OS Homo sapiens.
XX
PN WO200157271-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00662.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-496933/54.
XX
PT New spatially-addressable set of single exon nucleic acid probes,
PT useful for measuring gene expression in sample derived from human
PT breast, comprises number of single exon nucleic acid probes -
XX
XX Claim 27; SEQ ID NO 11100; 327pp + sequence listing; English.
XX
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting
CC the probes with a collection of detectably labelled nucleic acids
CC derived from mRNA of human breast, and then measuring the label
CC bound to each probe of the microarray. The probes are useful for
CC verifying the expression of regions of genomic DNA predicted to
CC encode proteins. They are useful for gene discovery, and for
CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater
CC diversity of probes for measuring gene expression, with far less bias
CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence. The
CC present sequence is a peptide encoded by a single exon nucleic acid
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 70 AA;
SQ
Query Match 14.1%; Score 365; DB 22; Length 70;
Best Local Similarity 100.0%; Pred. No. 8.5e-24;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 312 MIDTEFAKQTSLDVAQAQVVDVRVKRIHSDTFASGGERARFCPRHEDMTLLVRNFGYPLGE 371
Db 1 MIDTEFAKQTSLDVAQAQVVDVRVKRIHSDTFASGGERARFCPRHEDMTLLVRNFGYPLGE 60

QY 372 MSQPTSPAP 381
Db 61 MSQPTSPAP 70

RESULT 10
ABB33307
ID ABB33307 standard; Peptide; 70 AA.
XX
AC ABB33307;
XX
DT 04-FEB-2002 (first entry)
XX

Query Match 14.1%; Score 365; DB 22; Length 70;
Best Local Similarity 100.0%; Pred. No. 8.5e-24;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 312 MIDTEFAKQTSLDVAQAQVVDVRVKRIHSDTFASGGERARFCPRHEDMTLLVRNFGYPLGE 371
Db 1 MIDTEFAKQTSLDVAQAQVVDVRVKRIHSDTFASGGERARFCPRHEDMTLLVRNFGYPLGE 60

QY 372 MSQPTSPAP 381
Db 61 MSQPTSPAP 70

RESULT 11
ABB18767
ID ABB18767 standard; Protein; 70 AA.
XX
AC ABB18767;
XX
DT 23-JAN-2002 (first entry)
XX
DE Protein #766 encoded by probe for measuring heart cell gene expression.
KW Human; gene expression; heart; microarray; vascular system;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.

```

```

DE Peptide #813 encoded by human foetal liver single exon probe.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00669.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human fetal liver -
XX
XX Claim 27; SEQ ID NO 25942; 639pp + sequence listing; English.
XX
XX The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC foetal liver. The present sequence is a peptide encoded by a single exon
CC nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 70 AA;
SQ
Query Match 14.1%; Score 365; DB 22; Length 70;
Best Local Similarity 100.0%; Pred. No. 8.5e-24;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 312 MIDTEFAKQTSLDVAQAQVVDVRVKRIHSDTFASGGERARFCPRHEDMTLLVRNFGYPLGE 371
Db 1 MIDTEFAKQTSLDVAQAQVVDVRVKRIHSDTFASGGERARFCPRHEDMTLLVRNFGYPLGE 60

QY 372 MSQPTSPAP 381
Db 61 MSQPTSPAP 70

RESULT 11
ABB18767
ID ABB18767 standard; Protein; 70 AA.
XX
AC ABB18767;
XX
DT 23-JAN-2002 (first entry)
XX
DE Protein #766 encoded by probe for measuring heart cell gene expression.
KW Human; gene expression; heart; microarray; vascular system;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.

```

```
XX 09-AUG-2001.
PD 30-JAN-2001; 2001WO-US00666.
XX
XX 30-JAN-2001; 2001WO-US00666.
XX
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488899/53.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX hearts -
XX
XX Claim 15; SEQ ID No 20537; 530pp; English.
XX
XX The present invention relates to single exon nucleic acid probes for
XX measuring human gene expression in a sample derived from human heart (see
XX CCA21535-ABA41305). The present sequence is a protein encoded by one such
XX probe. The probes may be used for predicting, measuring and displaying
XX gene expression in samples derived from the human heart via microarrays.
XX By measuring gene expression, the probes are useful for predicting,
XX diagnosing, grading, staging, monitoring and prognosing diseases of the
XX human heart and vascular system e.g. cardiovascular disease,
XX hypertension, cardiac arrhythmias and congenital heart disease.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 70 AA;
SQ
Query Match 14.1%; Score 365; DB 22; Length 70;
Best Local Similarity 100.0%; Pred. No. 8.5e-24;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 312 MIDTEFAKQTSLDVAQAQAVVDRVKRIHSDTFASGGERARFCPRHEDMTLLVNFQYPLGE 371
DB 1 MIDTEFAKQTSLDVAQAQAVVDRVKRIHSDTFASGGERARFCPRHEDMTLLVNFQYPLGE 60
QY 372 MSQPTPSPAP 381
DB 61 MSQPTPSPAP 70
RESULT 12
AAM54096
ID AAM54096 standard; Protein; 70 AA.
XX
XX AAM54096;
AC
XX
XX 05-NOV-2001 (first entry)
DT
XX
XX Human brain expressed single exon probe encoded protein SEQ ID NO: 26201.
DE
XX
XX Human; brain expressed exon; gene expression analysis; probe;
XX microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
XX epilepsy; cancer.
XX
XX Homo sapiens.
OS
XX
XX WO200157275-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US00667.
PF
```

```
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains -
XX
XX Example 4; SEQ ID NO: 26201; 650pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is a protein encoded by one of
XX the probes of the invention.
XX
XX Sequence 70 AA;
SQ
Query Match 14.1%; Score 365; DB 22; Length 70;
Best Local Similarity 100.0%; Pred. No. 8.5e-24;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 312 MIDTEFAKQTSLDVAQAQAVVDRVKRIHSDTFASGGERARFCPRHEDMTLLVNFQYPLGE 371
DB 1 MIDTEFAKQTSLDVAQAQAVVDRVKRIHSDTFASGGERARFCPRHEDMTLLVNFQYPLGE 60
QY 372 MSQPTPSPAP 381
DB 61 MSQPTPSPAP 70
RESULT 13
AAM66486
ID AAM66486 standard; Protein; 70 AA.
XX
XX AAM66486;
AC
XX
XX 06-NOV-2001 (first entry)
DT
XX
XX Human bone marrow expressed probe encoded protein SEQ ID NO: 26792.
DE
XX
XX Human; bone marrow expressed exon; gene expression analysis; probe;
XX microarray; cancer; leukemia; lymphoma; myeloma.
XX
XX Homo sapiens.
OS
XX
XX WO200157276-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US00668.
PF
XX
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
```

XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488900/53.
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
XX Example 4; SEQ ID NO: 26792; 658pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention.
XX
XX Sequence 70 AA;
SQ
Query Match 14.1%; Score 365; DB 22; Length 70;
Best Local Similarity 100.0%; Pred. No. 8.5e-24;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 312 MIDTEFAKQTSLDAVAQAVVDRVKRIHSDTFASGGERARFCPRHEDMTLLVRNFGYPLGE 371
Db 1 MIDTEFAKQTSLDAVAQAVVDRVKRIHSDTFASGGERARFCPRHEDMTLLVRNFGYPLGE 60
QY 372 MSQPTPSPAP 381
Db 61 MSQPTPSPAP 70
RESULT 14
AAM14359
ID AAM14359 standard; Protein; 70 AA.
XX
AC AAM14359;
XX
XX 12-OCT-2001 (first entry)
XX
XX Peptide #793 encoded by probe for measuring cervical gene expression.
XX
XX Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer.
XX
XX Homo sapiens.
XX
XX WO200157278-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00670.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488901/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -
XX
XX Claim 27; SEQ ID No 19185; 487pp; English.
XX
XX The present invention relates to human single exon nucleic acid probes

CC (SENP: see AA110068-AA128459). The present sequence is a peptide encoded
CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs
CC can be used to produce a single exon microarray, which can be used for
CC measuring human gene expression in a sample derived from human cervical
CC epithelial cells. By measuring gene expression, the probes are therefore
CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 70 AA;
SQ
Query Match 14.1%; Score 365; DB 22; Length 70;
Best Local Similarity 100.0%; Pred. No. 8.5e-24;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 312 MIDTEFAKQTSLDAVAQAVVDRVKRIHSDTFASGGERARFCPRHEDMTLLVRNFGYPLGE 371
Db 1 MIDTEFAKQTSLDAVAQAVVDRVKRIHSDTFASGGERARFCPRHEDMTLLVRNFGYPLGE 60
QY 372 MSQPTPSPAP 381
Db 61 MSQPTPSPAP 70
RESULT 15
AAM26772
ID AAM26772 standard; Protein; 70 AA.
XX
AC AAM26772;
XX
XX 17-OCT-2001 (first entry)
XX
XX Peptide #809 encoded by probe for measuring placental gene expression.
XX
XX Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder.
XX
XX Homo sapiens.
XX
XX WO200157272-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00663.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488897/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta -
XX
XX Claim 27; SEQ ID No 27041; 654pp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENP:
CC see AA131315-AA157546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders.
XX

SQ Sequence 70 AA;

Query Match

14.1%; Score 365; DB 22; Length 70;

Best Local Similarity 100.0%; Pred. No. 8.5e-24;

Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 312 MIDTEPAKQTSIDAVAQAVVDVKRIHSDTFASGGERARFCPRHEDMTLLVNFQYPLGE 371

Db 1 MIDTEPAKQTSIDAVAQAVVDVKRIHSDTFASGGERARFCPRHEDMTLLVNFQYPLGE 60

QY 372 MSOPTPSPAP 381

Db 61 MSOPTPSPAP 70

Search completed: December 9, 2002, 22:47:18
Job time : 55.0526 secs

QY 9 LQSEQQSWTDDLPCHLSGVGSASNRYS---ADKGTESSHPPEDSWLKFRSENN--CF 63
Db 8 LYSNQPO-TVEAP---ASGGGLSQNGKFSYGYASSAGKRSMEDFFETRIDGINGEIVG 63
QY 64 LYGVFNQYDGNRVNFAQRLSAELLQLQNAEHAEE--ADVRVLLQAFDVVERSFLFSI 121
Db 64 LFGVFDGHHGARAAYVVRHLFSNLI-----THPKFISDTKSAITDAYNHTDSELLKSE 117
QY 122 DDLAEKASLOSQLPQGVPOHQLPQYQKILERLKTLEIRETSGGAMAVAVLLNNKLYVA 181
Db 118 N-----SHNRDAGSTASTAILVGDRLVVA 141
QY 182 NVGTRALLCKSTVDGLQVTLQNLVDHTTENEDELFRLSQLDAGKIKQVGIICGQESTR 241
Db 142 NVGDSRAVISR---GGKAIASVRDHKPDQSDERERENAG---GFVWAGVLA---VSR 191
QY 242 RIGDYKVKYGYTDIDLLSAAKSKPIIAEPEIHGAQPLDGVGTGLVLMSEGLYKALENAHG 301
Db 192 AFGORLUK-----QYVADPEIQ-BEKIDDTLEFLILASDGLWDF----- 231
QY 302 PQQANQEIAMI---DTEFAKQTSLDAAVAVVDRVKRIHSOTTFASGGERARFCPRHED 357
Db 232 ---SNEAAVAMVKEVDEP-----DSAKLVGEAIKRGSA-----N 265
QY 358 MTLVVRNFGYPLGEMSOPTSPAPAGGRVVPVSVYSSAQSTKTSVTLSLVNPSQGM 417
Db 266 ITCVVVRF-----LEKKSASSSHISSSSSKEAKEMPPLGDL 301
QY 418 VNGAHSASTLD---EATPILTINOSPTLTLOSTNTHTQSSSSSD 458
Db 302 AISSNEAKQVOIGNGKPNVTRKPDTSRSTDTLTLERNSTVD 346
RESULT 5
T05680
hypothetical protein F20M13.80 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 20-Jun-2000
C:Accession: T05680
R:Bevan, M.; Wedler, H.; Kutzner, M.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.
submitted to the Protein Sequence Database, February 1999
A:Reference number: Z15420
A:Accession: T05680
A:Molecule type: DNA
A:Residues: 1-395 <BEV>
A:Cross-references: EMBL:AL035540
A:Experimental source: cultivar Columbia; BAC clone F20M13
C:Genetics:
A:Map position: 4
A:Introns: 102/1; 224/3; 303/3
A:Note: F20M13.80
C:Superfamily: Arabidopsis thaliana hypothetical protein F20M13.80

Query Match 7.4%; Score 190.5; DB 2; Length 395;
Best Local Similarity 24.8%; Pred. No. 7.6e-06;
Matches 104; Conservative 60; Mismatches 150; Indels 111; Gaps 20;
QY 4 QRRSLQEQQPSWTDLDPLCHLSGVGSASNRYSADGKTESHPEDSWLKFRSENNCF 63
Db 49 QANSLLEDQSQ-----LESGLSHSDSGPFT----- 75

QY 64 LYGVFNQYDGNRVNFAQRLSAELLQLQNAEHAADVRVLLQAFDVVERSFLFSID 123
Db 76 FVGVDYDGHGGPETSFRFINDHMFHH--LKRFTEAQ-QCMSSEVIKKAFAQTEEGFLSIVTN 132
QY 124 ALAEKASLOSQLPQGVPOHQLPQYQKILERLKTLEIRETSGGAMAVAVLLNNKLYVANV 183
Db 133 -----QFOTRPOIATV-----GSCCLSVICDGKLYVANA 162
QY 184 GTNRALLCK-STVDG-LQVTLQNLVDHTTENEDELFRLSQLDAGKI-----KQVGI 233
Db 163 GDSRAVLQGMVRVTGEHAHTQLSAEHNASIESVRRELQALHPDHPDIIVLKNNVMRVKGI 222

QY 234 ICQESTRRIGDYKVKYG-----YTDIDLLSAAKSKPII-AEPEI--HGAQPLDGVGTG 283
Db 223 I---QVRSIGDYLKRSEFNREPIYAKFRLSPF-SKPLLSAEPAITVHTLPHD---Q 275
QY 284 FLVLMSEGLYKALEAAGHGPQANQEIAMIDTE----FAKQTSLDAAVAVVDRVKRIHS 339
Db 276 FIICASDGLWEHM-----SNQEAVIDVQNHPRNGIAKRLVKVALQEAACKREMYSD 327
QY 340 DTFASGGERARFCPRHEDMTLLVRNFGYPLGEMSOPTSPAPAGGRVVPVSVYSS-AQ 398
Db 328 LKKIDRGVRRHF---HDDITVIVVFFDTNL--VSRGSMRLRGFAVSVRGAGVNLPHNTLAP 382
QY 399 STSKT 403
Db 383 CITPT 387

RESULT 6

A55804
phosphoprotein phosphatase (EC 3.1.3.16) 2c, membrane-bound - Paramesidium tetraurelia
C:Species: Paramesidium tetraurelia
C:Date: 23-Mar-1995 #sequence_revision 05-Apr-1995 #text_change 07-Dec-1999
C:Accession: A55804
R:Klumpp, S.; Hanke, C.; Donella-Deana, A.; Beyer, A.; Kellner, R.; Pinna, L.A.; Schultz, J. Biol. Chem. 269, 32774-32780, 1994
A:Title: A membrane-bound protein phosphatase type 2C from Paramesidium tetraurelia. Purified from Paramesidium tetraurelia.
A:Reference number: A55804; MUID:95105156; PMID:7806499
A:Accession: A55804
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-300 <KLU>
A:Cross-references: GB:Z36985; NID:G537421; PID:el192609; PID:G2654382
C:Genetics:
A:Genetic code: SGC5
C:Keywords: phosphoric monoester hydrolase

Query Match 7.4%; Score 190; DB 2; Length 300;
Best Local Similarity 24.5%; Pred. No. 5.3e-06;
Matches 66; Conservative 51; Mismatches 104; Indels 48; Gaps 9;
QY 64 LYGVFNQYDGNRVNFAQRLSAELLQLQNAEHAADVRVLLQAFDVVERSFLFSID 123
Db 52 VFGVFDGHHGREGVAQFVEKHFDVDELLKNK-----NFKEQKFEEALKE 93
QY 124 ALAEKASLOSQLPQGVPOHQLPQYQKILERLKTLEIREIS--GGAMAVAVLLNNKLYVAN 182
Db 94 TPLKMDL-----LLTPGQKELNQYKATDTDESAGCTANVALIYKNTLYVAN 142
QY 183 VGTNRALLCKSTVDGLQVTLQNLVDHTTENEDELFRLSQLDAGKIKQVGIICQESTRR 242
Db 143 AGDSRSLCRNTN---HMSVDHKKPDNPEKSRIERAG---GFVSDGRVNGNLNLSRA 195

QY 243 IGDYKVKYGYTDIDLLSAAKSKPIIAEPEIHGAQ--PLDGVTFGLVLMSEGLYKALEAAH 300
Db 196 LGDLEYKR-----DNKLRNQELIILALPDVKKTBELTPQD---KFLMGCDGVFTLNHOE 247
QY 301 GPGQANQEI-AMIDTEFAKQTSLDAAVQ 328
Db 248 LLKQVNSTIGQAQVTELLKKAEDLDDQ 276

RESULT 7

P86355
T16E15.10 protein - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 27-Nov-2001
C:Accession: P86355
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.; Ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luos, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, C.J.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: F86355
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-281 <STO>
 A:Cross-references: GB:AE005172; NID:g9392666; PIDN:AAF87263.1; GSPDB:GN00141
 C:Genetics:
 A:Map position: 1

Query Match 7.3%; Score 188.5; DB 2; Length 281;
 Best Local Similarity 24.8%; Pred. No. 6,1e-06;
 Matches 83; Conservative 59; Mismatches 88; Indels 105; Gaps 17;

QY 30 GSASNR---YSADG-----KGTSEHPED-----SWLKFSENNCFIYGVFNQYDG 73
 DB 16 GSSSRSGKGRSDEGMIXYGSFLVKGNHPMEDYHVFNFNIO-DHDLGLFAIYDGHMG 74
 QY 74 NRVTFVAGRLSAELLGQLNAEHAE--ADVRRVLLQAFDVERSFLESIDALAEKASL 131
 DB 75 DSVPAVLQKRLFSNLT-----KEGEFVWDPRRSIAKAYEKTDAILNSD----- 120
 QY 132 QSLPEGVPOHQLPPQYQKILERLKTLEREISGAMAVAVALLN-NKLYVANVGTNRALL 190
 DB 121 -----LGR-----GSGTAVAILINRGKIMIANVGDSSRAVL 151
 QY 191 CKSTVDGLVQLNVNDH--TTEN---EDBLFRLSQLGDAGKIKVGIICQESRRIRGD 245
 DB 152 S-----HGAITQMSDHEPRTERSSIEDRGFVSNLPGVPRVN-----GQ 193
 QY 246 YVKKVGYTDIDLSAAKSKPIIAEPEIHQAOPLDGVGTGLVMSSEGLYK-----ALBAA 299
 DB 194 LAVSRAPGDKGLKTHLSSEPDIKEATV-----DSQTDVLLASDGIMKWMNEAMEIA 247
 QY 300 H--GFGANOIEIAMIDTEFAKQTSLDVAQAAYV 331
 DB 248 RRVKDPQAKAKELTA---EALRREKSDIDISCVV 278

RESULT 8

T06308
 protein phosphatase 2C homolog F11C18.60 - *Arabidopsis thaliana*
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)
 C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 15-Jun-2001
 C:Accession: T06308
 R:Bevan, M.; Terry, N.; Ardiles, W.; Buyschaert, C.; Dasseville, R.; De Clerck, R.; De ewes, H.W.; Mayer, K.F.X.; Schueller, C.
 submitted to the Protein Sequence Database, April 1999
 A:Reference number: Z15589
 A:Accession: T06308
 A:Molecule type: DNA
 A:Residues: 1-357 <BEV>
 A:Cross-references: EMBL:AL049607; GSPDB:GN00062; ATSP:F11C18.60
 A:Experimental source: cultivar Columbia; BAC clone F11C18
 C:Genetics:
 A:Gene: ATSP:F11C18.60
 A:Map position: 4
 A:introns: 39/3; 61/1; 97/2; 148/3; 190/3; 232/1; 257/3; 275/2; 293/3
 C:Superfamily: human phosphoprotein phosphatase 1A

Query Match 7.3%; Score 188.5; DB 2; Length 357;
 Best Local Similarity 24.6%; Pred. No. 9e-06;
 Matches 85; Conservative 65; Mismatches 145; Indels 51; Gaps 14;

QY 59 ENNCFLYGVNGYDGNRVTFVAGRLSAELLGQLNAEHAEADVRRVLLQAF---DVV-- 113
 DB 48 DNTSFL-GVYDGHGKIVSKFKCAKYLHQVLSD---AVAAGDVGTSLQKAFRRMDMMQ 103
 QY 114 -ERSFLESIDALAEKASLSQSLPEGV-----PQHQLPPOYQKILERLKTLEREISGA 166

DB 104 GQRGRBELA--VLGKINKFSGMIEGLIMSPRSGSANKPRDAMAFEEBPHDFAGPNSGS 161
 QY 167 MAVAVALLNNKLYVANVGTNRALLCKSTVYDGLQVQLNVNDHTTENEDELFRISQLGDAG 226
 DB 162 TCAVAVRDQKQFVNAVAGDSRCVSRKN---QAVNLSRDHKFDLEAKERI---LKAG 213
 QY 227 KIKVQVILIQG-ESTRRIGDYKVKGYTDIDLSAAKSKPIIAEPEIHQAOPLDGVTFGL 285
 DB 214 GFTHAGRVNGSINLSRALGDMEFKQ---NKLPEBKQIVTASPDVNVTELCD-DDFL 267
 QY 286 VLMSEGLYKALBAHGPQANOIEIAMIDTEFAKQTSLDVAQAAYDVRYKIRHSPTFASG 345
 DB 268 VLAQDGINQCM-----TSQQLVDFIHEQLNSETKLSVCEKTLDCLAQNT---SG 315
 QY 346 GERARFCPRHEDTLLVRNFGYPLGEMSOPTSPAPPAAGRVYVP 391
 DB 316 GEGC-----DNMTWILVRFKNPTSETELKPEASQAEQNHDEPSS 355

RESULT 9

S61650
 hypothetical protein YOR090C - yeast (*Saccharomyces cerevisiae*)
 N:Alternate names: hypothetical protein O3157; hypothetical protein YOR3157c
 C:Species: *Saccharomyces cerevisiae*
 C>Date: 09-Mar-1996 #sequence_revision 12-Apr-1996 #text_change 19-Apr-2002
 C:Accession: S61650; S66975
 R:Benes, V.; Andrade, M.A.; Rechmann, S.; Teodoru, C.; Banrevi, A.; Sander, C.; Valencia submitted to the EMBL Data Library, December 1995
 A:Description: Nucleotide sequence and analysis of a 130 kb fragment of yeast chromosome
 A:Reference number: S61643
 A:Accession: S61650
 A:Molecule type: DNA
 A:Residues: 1-572 <BEN>
 A:Cross-references: EMBL:X94335; NID:g1262139; PID:e217726; PID:g1164936
 R:Voss, H.; Benes, V.; Rechmann, S.; Teodoru, C.; Schwaiger, C.; Paces, V.; Ansoerge, W. submitted to the Protein Sequence Database, July 1996
 A:Reference number: S66965
 A:Accession: S66975
 A:Molecule type: DNA
 A:Residues: 1-572 <VOSS>
 A:Cross-references: EMBL:Z74998; NID:g1420258; PID:e252349; PID:g1420259; MIPS:YOR090C
 A:Experimental source: strain S288C
 C:Genetics:
 A:Gene: SGD:PTCS
 A:Cross-references: SGD:S0005616
 A:Map position: 15R

Query Match 7.1%; Score 183.5; DB 2; Length 572;
 Best Local Similarity 22.9%; Pred. No. 4.3e-05;
 Matches 107; Conservative 74; Mismatches 159; Indels 127; Gaps 21;

QY 46 SHPPEDSMWK-----FRSEN-----NCFLYGVNGYDGNRVTFVAGRLSAELL----- 89
 DB 162 NHPFIEDDHVEQIITPIESBDGKSIEKDYFFGFIPOGHG---PTSEKLSKDLVRVYA 217
 QY 90 --IGQL---NAEHAEADVRRVLLQAFDVERSFLESIDALAEKASLSQSLPQGVPOHQL 144
 DB 218 YQLQGVYDQNKTVFHSNPQLIDSA---ISKGLKLDNDLVISFKLRQ----- 264
 QY 145 PPQYQKILERLKTLEREISGAMAVAVL--LNKLYVANVGTNRALLCKSTVDG-LQYT 201
 DB 265 DPNTNTNANTLPAI-----SGSCLLSLVYSTNSILKVANVTGSRALLIGLDNREGMWTYK 319
 QY 202 QLVNDHTTENEDELFRISQLGDAGKIKVGIICQ-ESTRRIGDYKVKGYTD-----ID 256
 DB 320 SLSTDQGDNDLDEVRRIRKEHPGPVIRNGRILGSLQPSRAGDYRYKIKVEYDKPLSD 379
 QY 257 LLSAA-----SKSP-IIAEPEIHQAOPLDGVGTGLVMSSEGLYKALBAHGPQ 304
 DB 380 LPEVAKLIFRRRPRDFKTPPVVTAPEVITSK-IGENTKRMVWGSQDLFELL----- 430
 QY 305 ANQEIAMIDTEFAKQTSL-----DAVAQAAYDVRYKIRHSPTFASG 346

C:Genetics:
A:Gene: At2g34740, T29F13.5
A:Map position: 2
A:Introns: 46/3; 139/1; 190/3

Query Match 6.9%; Score 177.5; DB 2; Length 239;
Best Local Similarity 25.9%; Pred. No. 2.9e-05;
Matches 78; Conservative 50; Mismatches 90; Indels 83; Gaps 14;

Qy 39 ADGKTESHPEDSWLKFRSENNCFLYGFNGYDGNRVTFVAQRLSAELLQLGNAEHA 98
Db 7 ADTKTVKGH-----NLGIYAFDGHSGSDVADYQLQNLPNIL-----S 45
Qy 99 EADVRRVLLQAFDVERSFLESIDDLAEKASLQSQLPBGVQHQLPPOYQKILERLKT 158
Db 46 QPFDKFNPKKA---IKRAY-KSTDYIYL-----QNVGPR----- 76
Qy 159 EREISGAMAVAVALLN-NKLVANVGTNRALLCKSTVDCLOVQTQNTDHTTENEDELFR 217
Db 77 ----GGSTAVTAIVDGGKIVANVAGDSRALICR---ESDVAKQITVDHEPKERDLVK 128
Qy 218 LSQGLDACKIKQVGIICQO-ESTRIRIGDYKVGYYTIDILLSAAKSPITAEPIHGAQ 276
Db 129 -SKGGSVSKPQGNPRVVDQQLAMTRAFSGDGLKEHISVI-----PNIETAIHND- 177
Qy 277 PLDGVTFGLVLMSEGLYKAL-----EAAHGPQANQETIAMITPEFAKQTSLDVAQAV 330
Db 178 ----TKFPLILASDGLMKVMSNDEVDQIKKRGNAEAKMLIDKALAR-GSKDIDISCV 231
Qy 331 V 331
Db 232 V 232

RESULT 13
S55457
phosphoprotein phosphatase (EC 3.1.3.16) 2C - Arabidopsis thaliana
N:Alternate names: protein phosphatase 2C
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 01-Mar-2002
C:Accession: S55457; S53581
R:Yamamoto, M.
submitted to the EMBL Data Library, August 1994
A:Reference number: S55457
A:Accession: S55457
A:Molecule type: mRNA
A:Residues: 1-399 <YAM>
A:Cross-references: EMBL:D38109; NID:G633027; PID:d1007865; PID:G633028
R:Kutromoti, T.; Yamamoto, M.
Nucleic Acids Res. 22, 5296-5301, 1994
A:Title: Cloning of cDNAs from Arabidopsis thaliana that encode putative protein phosphatase
A:Reference number: S53581; MUID:95116318; PMID:7816619
A:Accession: S53581
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 82-399 <KUR>
A:Cross-references: GB:D38109
C:Superfamily: Arabidopsis thaliana hypothetical protein F7A7.220
C:Keywords: phosphoric monoester hydrolase

Query Match 6.8%; Score 175; DB 2; Length 399;
Best Local Similarity 25.5%; Pred. No. 9.8e-05;
Matches 93; Conservative 58; Mismatches 117; Indels 96; Gaps 20;
Qy 7 SLLOSEQSQSWTDDLPYLCHLSGVGSASNSYSADGKTES--HPEDSWLKFRSENNCFLY 65
Db 91 SYTEBA---SFSQVDP-----KIGTTSVCGRRDMDVASHIP---SLQNRSENNH--Y 138
Qy 66 GVFNGYDGNRVTFVAQRL-----SAELLQLGNAEHAADVRRVLLQAFDVERSFLE 119
Db 139 GVFDDGSHVAEKCRERLHDIVKKEVEVMASD---EMTETWVKSFQKMDKEVSQRECNL 195
Qy 120 SLDDAL-AEKASLQSQLPBGVQHQLPPOYQKILERLKTLEKEISGAMAVAVAVALLNKL 178

Db 196 VVNGATRSKMNKSCRCIQS-----POCDA-----GSTAVVSVTPPEKI 234
Qy 179 YVANVGTNRALLCKSTVDCLOVQTQNTDHTTENEDELFRLSQGLDAGKI-----KQVG 232
Db 235 IVNCGDSRAVLGRNV-----AIPLSVDHKPRPBDLIRIQAG---GNVIWDGARVLG 287
Qy 233 IICQESTRIRIGYKVGYYTIDILLSAAKSPITAEPIHGAQPLDGVTFGLVLMSEGL 292
Db 288 VLA---MSRAIGDNYLK-----PYVIPDPEVTWTDRTD-EDECLILASDGL 329
Qy 293 YKAL--EAA-----HGPQANQETIA-----MIDTEFAKQTSLDVAQAVVD 332
Db 330 MDVAVNETACGAVMCLRGAGGDSDAHNAACSDALLITKALAKRSS-DNVSVVVD 388
Qy 333 RVKR 336
Db 389 LRKR 392

RESULT 14
T47644
protein phosphatase 2C-like protein - Arabidopsis thaliana
N:Alternate names: protein T15C9.50
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 21-Jul-2000
C:Accession: T47644
R:Mewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.
submitted to the Protein Sequence Database, April 2000
A:Reference number: Z24470
A:Accession: T47644
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-409 <MEM>
A:Cross-references: EMBL:AL132970
A:Experimental source: cultivar Columbia; BAC clone T15C9
C:Genetics:
A:Map position: 3
A:Introns: 46/1; 336/3
A:Note: T15C9.50
C:Superfamily: Arabidopsis thaliana hypothetical protein P20M13.80

Query Match 6.8%; Score 175; DB 2; Length 409;
Best Local Similarity 25.2%; Pred. No. 0.0001;
Matches 101; Conservative 56; Mismatches 132; Indels 112; Gaps 21;

Qy 25 HLSG-----VGSASNRSYSADGKTESHP-----PEDSWLKFRSENNCFLYGFNGYD 72
Db 71 HITGEPSMAVVOANN--LLEDHSGLESGLPSLHESGPEATFV-----GYIDGHG 117
Qy 73 GNRVTNFVAQRLSAELLQLGNAEHAADVRRVLLQAFDVERSFLESIDDLAEKASIQ 132
Db 118 GPEAARFVNDRLF-----YNIKRYTSBQGM---SPDVIIRGFATVEEFL----- 160
Qy 133 SOLPEGVQHQLPPOYQKILERLKTLEKEISGAMAVAVALLNKLKYVANVGTNRALLCK 192
Db 161 ----GLVQ-----EQWKTPQIASVGACCLVIGVCGLLYVNAADSDRVVLCK 204
Qy 193 --STVDCLOVQTQNTDHTTENEDELFRLSQGLDAGKI-----KQVGIIICQESTR 242
Db 205 VANPFEKELKAVQLSTEHNSIBSVREELRLHPDDNIVVLGHKVKRVGII---QVSHS 261
Qy 243 IGD-YKVGYYTIDILLSAAK-----SKPIT-AEPEI--HGAQPLDGVTFGLVLMSEGLY 293
Db 262 IGDVYLKRAAFNGEPRLPKFRVPERFEKPIMAEPTITVHKIHPED---QFLIFASDGLM 318
Qy 294 KALEAHHGPQANQETIAMITD---EFAKQTSLDVAQAVDVKYKIHSDTPASGGERA 349
Db 319 EHL-----SNQEAVDIVNSCPRNGVARKLVKALQAAKREWRYSDELKIERGIR 370
Qy 350 RFPCEHEDMTLLV-----RNFGYPLGEMSQPTSPAPAGG 385
Db 371 HF--HDDITIVIVVFLHATNFA-----TRTPISTVKG 399

Job time : 31.8532 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 9, 2002, 22:45:29 ; Search time 54.4488 Seconds
(without alignments)
1907.257 Million cell updates/sec

Title: US-09-830-144-4
Perfect score: 2580
Sequence: 1 MAQRRLSQEQPQSWTDD.....AEFYRLSWDHGQSVVTAP 504

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 21:.*
1: sp_archaea:.*
2: sp_bacteria:.*
3: sp_fungi:.*
4: sp_human:.*
5: sp_invertebrate:.*
6: sp_mammal:.*
7: sp_mmc:.*
8: sp_organelle:.*
9: sp_phase:.*
10: sp_plant:.*
11: sp_rodent:.*
12: sp_virus:.*
13: sp_vertebrate:.*
14: sp_unclassified:.*
15: sp_rvirus:.*
16: sp_bacteriap:.*
17: sp_archaeap:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2512	97.4	500	11 Q8R0D1	O8r0d1 mus musculus
2	1924	74.6	498	13 Q73614	O73614 xenopus lae
3	275	10.7	52	11 Q9CV62	Q9cv62 mus musculus
4	258.5	10.0	386	5 Q93375	Q93375 caenorhabdi
5	205.5	8.0	311	10 Q8RXV3	Q8rxv3 arabidopsis
6	205	7.9	319	10 Q8S8Z0	Q8s8z0 mesembryant
7	202	7.8	493	10 Q9LUS8	Q9lus8 arabidopsis
8	194.5	7.5	282	10 Q9S9Z7	Q9s9z7 arabidopsis
9	193.5	7.5	377	10 Q9LNF4	Q9lnf4 arabidopsis
10	192.5	7.5	355	10 Q81716	Q81716 arabidopsis
11	191.5	7.4	243	10 Q8S3P1	Q8s3p1 oryza sativ
12	191	7.4	816	10 Q9MLV6	Q9mlv6 arabidopsis
13	190.5	7.4	283	10 Q9AE3	Q9ae3 arabidopsis
14	190.5	7.4	348	10 Q9LEW5	Q9lew5 arabidopsis
15	190.5	7.4	395	10 Q9SZN2	Q9szn2 arabidopsis
16	190.5	7.4	400	10 Q8W4N8	Q8w4n8 arabidopsis

17	188.5	7.3	281	10 Q9LME4	Q9lme4 arabidopsis
18	188.5	7.3	357	10 Q9S253	Q9s253 arabidopsis
19	187.5	7.3	379	10 Q9LSN8	Q9lsn8 arabidopsis
20	186.5	7.2	385	10 Q9FKX4	Q9fkx4 arabidopsis
21	186	7.2	378	10 Q8S2S5	Q8s2s5 thellungiel
22	184.5	7.2	282	10 Q9ZPL8	Q9zpl8 lotus japon
23	184.5	7.2	420	10 Q8VZD9	Q8vzd9 arabidopsis
24	183.5	7.1	572	3 Q12511	Q12511 saccharomyC
25	183	7.1	305	10 Q942N4	Q942n4 oryza sativ
26	182.5	7.1	420	10 Q94AT1	Q94atl arabidopsis
27	182	7.1	271	10 Q9SMJ9	Q9smj9 sporobolus
28	181.5	7.0	238	10 Q9C9R2	Q9c9r2 arabidopsis
29	181	7.0	307	10 Q9FGM3	Q9fgm3 arabidopsis
30	180	7.0	355	10 Q9M8J1	Q9m8j1 arabidopsis
31	180	7.0	366	10 Q9FOY2	Q9foy2 zea mays (m
32	178.5	6.9	361	10 Q9LZ09	Q9lzo9 arabidopsis
33	178	6.9	392	4 Q9H0C8	Q9h0c8 homo sapien
34	177.5	6.9	239	10 Q64583	Q64583 arabidopsis
35	175	6.8	384	10 Q94CL8	Q94cl8 arabidopsis
36	175	6.8	409	10 Q9M2W1	Q9m2w1 arabidopsis
37	173.5	6.7	379	10 Q9SDI2	Q9sdi2 arabidopsis
38	172.5	6.7	528	10 Q93YS2	Q93ys2 arabidopsis
39	172	6.7	464	10 Q9FXE4	Q9fxe4 arabidopsis
40	169.5	6.6	372	4 Q96NT4	Q96nt4 homo sapien
41	169.5	6.6	392	11 Q9Z1Z6	Q9z1z6 rattus norv
42	169.5	6.6	416	10 Q9FEW0	Q9few0 nicotiana t
43	169	6.6	389	10 Q81773	Q81773 arabidopsis
44	169	6.6	416	10 Q9FLI3	Q9fl13 arabidopsis
45	168	6.5	423	10 Q9MLV8	Q9mlv8 arabidopsis

ALIGNMENTS

RESULT 1

Q8R0D1 PRELIMINARY; PRT; 500 AA.
ID Q8R0D1
AC Q8R0D1;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DE 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Similar to mitogen-activated protein kinase kinase 7
DE interacting protein 1 (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC027054; AAH27054.1; -.
KW Kinase.
FT NON TER 1
SQ SEQUENCE 500 AA; 54413 MW; 97529D2E2AE4DA34 CRC64;

Query Match 97.4%; Score 2512; DB 11; Length 500;
Best Local Similarity 97.2%; Pred. No. 4.2e-168;
Matches 488; Conservative 9; Mismatches 3; Indels 2; Gaps 1;

QY 3 AQRRLSQEQPQSWTDDLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKPFSENNC 62

Db 1 AQRRLSQEQPQSWTDDLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKPFSENNC 60

QY 63 FLYGVFNGYDGNRVTFNFFVAQRLSAELLQLQNAEHAEDVRRVLQAQDVVVERSFLSID 122

Db 61 FLYGVFNGYDGNRVTFNFFVAQRLSAELLQLQNAEHAEDVRRVLQAQDVVVERSFLSID 120

QY 123 DALAEKASLQSQLPQGVQHQHLPPOYQKILERLKTLEISGGAMAVVAVLLNNKLYVAN 182

Db 121 DALAEKASLQSQLPQGVQHQHLPPOYQKILERLKTLEISGGAMAVVAVLLNNKLYVAN 180

QY	183	VCNRALLCKSTVDGLQVQTNVHTHENEDELFRLSQLGLDAGKIKOVGIIIGGESTR	240
Db	181	VGTRRALCKSTVGLQVQTNMDHTTENDEBELFRLSQLGLDAGKIKOVGICGOSTR	240
QY	243	IGDVKKVKGTYDIDLLLSAAKSKPIIAEPELHGAQPLDVGTVLWSEGLYKALEAAGP	302
Db	241	IGDVKKVKGTYDIDLLLSAAKSKPIIAEPELHGAQPLDVGTVLWSEGLYKALEAAGP	300
QY	303	GOANOEIAAMIDTEPAKOTSIDAVAQAVDVRKRIHSDTPTASGGERARFCFRIEDMTLLV	362
Db	301	GOANOEIAAMIDTEPAKOTSIDAVAQAVDVRKRIHSDTPTASGGERARFCFRIEDMTLLV	360
QY	363	RNFQPLGEMSQPPTPPAPAGRVVYVSVYSSASTSKTSVTLSLWVMSQGMNGAH	422
Db	361	RNFQPLGEMSQPPTPPAPAGRVVYVSVYSSASTSKTSVTLSLWVMSQGMNGAH	418
QY	423	SASTIDEATPPLTNQSPPLTLQSTNTHTOSSSSSDGLFRSRPAHSLPBGEDGRVEPY	482
Db	419	SASTIDEATPPLTNQSPPLTLQSTNTHTOSSSSSDGLFRSRPAHSLPBGEDGRVEPY	478
QY	483	DFAEFYRLMSVDHGQSQSVTAP 504	
Db	479	DFAEFYRLMSVDHGQSQSVTAP 500	
RESULT 2			
ID	073614	PREDIMINARY;	PRT; 498 AA.
AC	073614;		
DT	01-AUG-1998 (TREMBlrel. 07, Created)		
DT	01-AUG-1998 (TREMBlrel. 07, Last sequence update)		
DT	01-MAR-2002 (TREMBlrel. 20, Last annotation update)		
DE	Tab1.		
OS	Xenopus laevis (African clawed frog).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;		
OC	Xenopodinae; Xenopus.		
OK	NCBI_TaxID=8355;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=98130593; PubMed=9463380;		
RA	Shibuya H., Iwata H., Masuyama N., Gotch Y., Yamaguchi K., Irie K.,		
RA	Matsumoto K., Nishida E., Ueno N.;		
RT	"Role of TAK1 and TAB1 in BMP signaling in early Xenopus		
RT	development."		
RL	EMBO J. 17:1019-1028 (1998).		
DR	EMBL; U92031; AAC14009.1; .		
DR	InterPro: IPR001932; PP2C-like.		
DR	Pfam; PF00481; PP2C; 1.		
DR	SMART; SM00332; PP2C; 1.		
SO	SEQUENCE 498 AA; 54450 MW; 654AB8DE2353524 CRC64;		
Query Match			
	74.6%; Score 1924; DB 13; Length 498;		
	Best Local Similarity 75.0%; Pred. No. 8e-127;		
	Matches 378; Conservative 54; Mismatches 64; Indels 8; Gaps		
QY	1	MAAQRSLLOSFQPSWTDLPLCHLSGVGSA SNRSYADCKGTESHPPEDSWLKFESEN	60
Db	1	MAAPRRNLHSG---SWTDLPLCLNLSGVGSA SNQYNSSEGLGDEHPEDNMWIKFEGDN	57
QY	61	NCFLYGVFGYGNKRVTFNVQORISAEILLQLMAHEHAQVRRVLLQAPFVERFLES	120
Db	58	NTLYGVFPYXEGTATSVFGRLAAEILLQLDQDPDVAEYHAKVLLQAPFVERFLES	117
QY	121	IDDAIAEKASLQSQPEGVPOHQLPPQYOKLLERLKTREISGGA MAVAVALLNKLTV	180
Db	118	IDDCIAEKASLQSLPEGLHOTLPSEQYOKIVDRNLITKEKIIYGGAWIVVLIYNSKLV	177
QY	181	ANVGNTRALLCKSTVDGLQVQTNVHTTENDEBELFRLSQLGLDAGKIKOVGIIIGGEST	240
Db	178	ANVGNTRALLCKSTVDGLQVQTNMDHTTENDEBELFRLSQLGLDTTKIKOVGVIGGOOST	237
QY	241	RKIGDYKKVKGTYDIDLLLSAAKSKPIIAEPELHGAQPLDVGTVGLVWSEGLYKALEAAH	300

```

Db      238 RRIIDYKVKVMFNNDIELLSTAKSKPTAEPIHCQPLDGVGTGLVLMSEGLYALLESBAH    297
Oy      301 GPGQANDGEIAMITTEPRKQTSLDVAQAAYVDRAKRLHSDFPAGGEPARCPRHEMDTL    360
Db      298 GPGQANDGEIAMIMITEFPAKVSLDEVAQALVERKRLIHHDPFAGGERAKYCASKHEMDTL    357
Oy      361 LVRRNGVPLGMSQPPTSPPAPAGARVYPVPVSSAQSSTKTSVTLSLVMPSCQGMWNG    420
Db      358 LVRNIGVPLVLOISPTLT--PTQGRLYPVSPVPSAONTSKTSTVLTLVMPSCQPMWNG    415
Oy      421 AHA$ASTIDEATPTLLTNOSPPTLLTQSTNTHTQSSSSSDGGLFRSRPAHSLPAGEDRVEP    480
Db      416 TNSSSTLDGGTSTL--QSPATLQSTNHTHTQSSSSSDGGLFRSRPLPSLPQDEDGRVEP    473
Oy      481 YVDAEAFRLM$SVDHGO-SVYA 503
Db      474 YVDFDFRYRLMNNAEHNDPTLLTA 497

RESULT 3
OyCv62          PRELIMINARY;          PRT;          52 AA.
O9CV62
AC              O9CV62;
DT      01-JUN-2001 (TREMBLrel. 17, Created)
DT      01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT      01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE      2310012M03RIk protein (Fragment).
GN      2310012M03RIK.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
RP      STRAIN=C57BL/6J; TISSUE=TONGUE;
RX      MEDLINE=21085660; PubMed=11217851;
RA      Kawai J., Shinagawa A., Shibara K., Yoshino M., Itoh M., Ishii Y.,
RA      Arakawa T., Hara A., Fukunishi Y., Komoh H., Adachi J., Fukuda S.,
RA      Aitawa K., Itawa M., Nishi K., Kiyosawa H., Kondo S., Yamanka I.,
RA      Sakata T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Satto R.,
RA      Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA      Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA      Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA      Schirml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Mashio T.,
RA      Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA      Blake J., Botfield D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA      Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA      Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
RA      Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momberees P.,
RA      Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA      Saeki H., Sato K., Schoenbach C., Seya T., Shibaya Y., Storch K.-F.,
RA      Suzuki H., Toyok-Oka K., Wang K.H., Weltz C., Whitlaker C., Wilming L.,
RA      Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsunki S.,
RA      Hayashizaki Y.;
RT      "Functional annotation of a full-length mouse cDNA collection.";
RL      Nature 409:685-690(2001).
DR      EMBL; AK009321; BAB86215.1..-
DR      MGD; WGI:1913763; 2310012M03RIK.
FT      NON TER                1
SQ      SEQUENCE   52 AA;  5712 MW;  813E29B1639920A6 CRC64;

Query Match      10.7%; Score 275; DB 11; Length 52;
Best Local Similarity 98.1%; Pred. No. 1.4e-12;
Matches      51; Conservative 1; Mismatches 0; Indels 0; Gaps 0

Oy      453 $SSSSDGGLFRSRPAHSLPAGEDRVEPVDAEAFRLM$SVDHGQSQSYTPA 504
Db      1     $SSSSDGGLFRSRPAHSLPAGEDRVEPVDAEAFRLM$SVDHGQSQSWTPA 52

```



```

DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Protein phosphatase 2C.
CN MPC9.
OS Mesembryanthemum crystallinum (Common ice plant).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Aizoaceae; Mesembryanthemum.
OX NCBI_TaxID=3544;
RN [1]
RP SEQUENCE FROM N.A.
RA Miyazaki S., Izumi S., Fukuhara T., Bohner H.J.;
RT "Plant Protein Phosphatases 2C? A large protein family serving
RT complex functions.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB083482; BAB88944.1; -
SQ SEQUENCE 319 AA; 34477 MW; 0BD228603E34CFB0 CRC64;

Query Match 7.9%; Score 205; DB 10; Length 319;
Best Local Similarity 23.1%; Pred. No. 1.9e-06;
Matches 83; Conservative 61; Mismatches 109; Indels 106; Gaps 14;

QY 12 EQQPSMTDDLPLCHLSGVSASNRYSADGKTESHPEDSWLKRSPNN-----CF 63
DB 17 BEQP-----VSGGIGSQNGKFS--YGVAASPGRKSSMEDPFTYTRIDGVEGEVVG 63
QY 64 LVGVNGYDGNVTNFWQRLSAEILLQLNAEHA--ADVRVLLQAFDVERSFILESI 121
DB 64 LFGVDDHGGAARAEVYKQNLFSNLI-----KHKPFSIDTKSAIAEAYTHTDSEFLKSE 117
QY 122 DDLAAEKASLSQSLPEGVPOHQLPPQYOKIERLKLTEREISGGAMAVAVLNNKLYVA 181
DB 118 N-----TQNRD--AGSTASTAILVGRLLVA 141
QY 182 NGNTRALLCKSTVDGLQVLTQNLVDHTTENEDLFRLSQLG---LDACKIKQVIGIQE 238
DB 142 NVGDSRAVTCR---GGEIAVSRDHKPKQSDERQRIEDAGSFVWAGTWKVGVLAA--- 194
QY 239 STRRIGDYVKVKGTYDIDLLSAKSKPILAEPEIHGAQPLDVGTVGLVMEGLYKALEA 298
DB 195 VSRAGCDKLK-----QYVADBEIQ--EEVVDSSLEFLLASDGLMDVVTN 239
QY 299 AHG-----PGQANQETIAMITFEAKQTSLDAVAQAVV-----DRYKRHSST 341
DB 240 EEAVTWVKPIQDTEAKAKLMQEAAYORGSADNITCVVAFRLDNPIDSSNRISNGHDT 298

RESULT 7
Q9LUS8 PRELIMINARY; PRT; 493 AA.
ID Q9LUS8;
AC Q9LUS8;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Similarity to protein phosphatase-2c.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=COLUMBIA;
RC STRAIN=COLUMBIA;
RX MEDLINE=20277480; PubMed=10819329;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT features of the regions of 4,504,864 bp covered by sixty P1 and TAC

```

```

RT clones.";
RL DNA Res. 7:131-135(2000).
DR EMBL; AB022217; BAB02747.1; -
DR InterPro; IPR001932; PP2C-like.
DR Pfam; PF00481; PP2C; 1.
DR SMART; SM00332; PP2C; 1.
DR SMART; SM00331; PP2C; SIG; 1.
SQ SEQUENCE 493 AA; 53613 MW; 613ED80B06B8C844 CRC64;

Query Match 7.8%; Score 202; DB 10; Length 493;
Best Local Similarity 23.5%; Pred. No. 6.2e-06;
Matches 96; Conservative 59; Mismatches 133; Indels 120; Gaps 18;

QY 24 CHLSGVSAS--NRYSADGKTE---SHPPEDSWLKRSENNCFLY--GVENGYDGNVT 77
DB 129 CYLTSSSPSSSVNEGFLSAMEVQVAGAGEDRVOAVCSFENGMWLCALYDGFNGRDAA 188
QY 78 NFVAQRL-----SAELLIGQLNAEHAED-----VR 103
DB 189 DFLACTLYESIYFHQLDLDRQMKQTKSDDDGKLELISNIVDYSSTDLPRQGLVDCIN 248
QY 104 RVLLQAFDVERSFILESIDDLAAEKASLSQSLPEGVPOHQLPPQYOKIERLKLTEREIS 163
DB 249 RALFQA---ETDFLMVGEQEMEERPD-----VS 274
QY 164 GGMAYVAVLNNKLYAVAVGTRNALLCKSTVDG---LQVLTQNLVDHTTENEDLFRLSQ 220
DB 275 VGSQVAVTLVLGKDLVYVLMIGSRAVL--ATYNGNKKLQAVQLTBDHTVDNVEEARLLS 332
QY 221 LGSADGKIKQVIGIQG--ESTRRIGDYKVKGYTYDIDLLSA-----KSPPIA- 268
DB 333 EHLDDPKIYVGKIKKGLKLVTRALG-----VGILKKEKNDALMGLRYRNLSSPPYVS 387
QY 269 EP--EIHGAQPLDVGTVGLVLMSEGLYKALEAHGEGQANQETIAMITFEAKQTSLD-- 324
DB 388 EPGMRVHKITESD--HFVIASDGLDFDF-----SNEBAIGLVHSFVSNPSGCPA 436
QY 325 -----AVQAQVYDVKRIHSDTFPAGSGERARFCPRHEMTLLVNFQ 366
DB 437 KLLERLVAAQAARAGFTLBEITNVAPGRRRY--HDDVTIMVITLGG 481

RESULT 8
Q9S9Z7 PRELIMINARY; PRT; 282 AA.
ID Q9S9Z7;
AC Q9S9Z7;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE P21H2.4 protein (protein phosphatase type 2C-like protein).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=CV. COLUMBIA;
RC Vysotskaya V.S., Schwartz J.R., Yu G., Toriumi M., Lenz C., Liu S.,
RA Lee J., Liu A., Li J., Kremenetskaia I., Luros J., Gonzalez A.,
RA Altafi H., Araujo R., Brooks S., Buehler E., Chao O., Conn L.,
RA Conway A.B., Dunn P., Hansen N., Huizer L., Khan S., Kim C., Palm C.,
RA Rowley D., Shim P., Walker M., Davis R.W., Ecker J.R.,
RA Federspiel N.A., Theologis A.;
RT "The sequence of BAC P21H2 from Arabidopsis thaliana chromosome 1.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=CV. COLUMBIA;
RC Theologis A.;
RA Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [3]

```



```
RP SEQUENCE FROM N.A.
RA Lam B., Southwick A., Karlin-Neumann G., Nguyen M., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Chen H., Cheuk R.,
RA Chung M.K., Kim C., Lin J., Liu S.X., Pham P.K., Sakano H., Shinn P.,
RA Yamada K., Ecker J., Theologis A., Davis R.W.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC007894; AAD46006.1; -
DR EMBL; AF370608; AAK43927.1; -
DR HSSP; P35813; IAGQ.
DR InterPro; IPR001932; PP2C-like.
DR Pfam; PF00481; PP2C; 1.
DR SMART; SM00332; PP2Cc; 1.
DR SMART; SM00331; PP2C-SIG; 1.
SQ SEQUENCE 282 AA; 30984 MW; 36404709FB03DA8D CRC64;

Query Match      7.5%; Score 194.5; DB 10; Length 282;
Best Local Similarity 25.4%; Pred. NO. 8.7e-06;
Matches 86; Conservative 45; Mismatches 97; Indels 111; Gaps 15;

QY 29 VCSASNRS-----YSAG-----KGTESHPPED-----SWLKFRSENNCFYGVFNQYD 72
Db 16 VGRASTSGKGRNNDGEIKFGVSLYKGRANHPMEDYHVSKEFKVI-DGNELGLFAIYDGH 74
QY 73 GNRVTNFVAQRLSAELLQLQNAEADVRVLLQAFDVVERSFLSIDDALAEKASLQ 132
Db 75 GERVPAIYQKHLFSNI-----LKEEFQRYDPQRSIIAAEYKTD----- 112
QY 133 SOLPEGVFQHQLPPOYQKILERLKLTEREISGGANAVAVALLN-NKLYVANVGNRALLC 191
Db 113 -----QAILSHSSDLGR---GGSTAVTATLNGRRLRVANVNGSRAVLS 153
QY 192 KSTVDGLQVTLQNVNHTTENEDELFRLSQLGDA-----GKIQVGIICQESTRIGD 245
Db 154 Q-----GGQAQMTIDHEPHT-----RLSTEGKGFVSNMGPVPRVN-----GQ 194
QY 246 YKVKYGYTDIDLSSAAKSPIIAEPIHGAQPLDGVTCFLVLMSEGLYKALEAAHGPQOA 305
Db 195 LAVSRAFGDKSLKTHLRSDPDVKDSSI-----DDHTDVLVSLASDGLWKVM-----A 240
QY 306 NOE-----IAAMIDTEPAKOTSLDAVAQAVV 331
Db 241 NQEAIDARIIRIKPLKAAKELTTEALRDSKDDISCIVV 279

RESULT 9
Q9LNF4
ID Q9LNF4 PRELIMINARY; PRT; 377 AA.
AC Q9LNF4
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DE 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE F21D18.27 (Protein phosphatase-2C, putative).
GN T2J15.5.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
[1]
RN SEQUENCE FROM N.A.
RA Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,
RA Shann P., Altafi H., Bei B., Chin C., Chioi J., Choi E., Conn L.,
RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
RA Ecker J.R.;
RT "Genomic sequence for Arabidopsis thaliana BAC F21D18 from chromosome I.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
```

```
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
[4]
RN SEQUENCE FROM N.A.
RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA Khan S., Kim C., Altafi H., Bei B., Chin C., Chioi J., Choi E.,
RA Conn L., Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B.,
RA Lee J., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
RA Theologis A., Ecker J.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
[5]
RN SEQUENCE FROM N.A.
RC STRAIN=CV; COLUMBIA;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Lib S.X., Liu Z.A., Luros J.S., Maiti R., Marziali A.,
RA Miltcher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzer S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis thaliana.";
RL Nature 408:816-820 (2000).
DR EMBL; AC023673; AAF79528.1; -
DR EMBL; AC051631; AAG51521.1; -
DR HSSP; P35813; IAGQ.
DR InterPro; IPR000222; PP2C.
DR InterPro; IPR001932; PP2C-like.
DR Pfam; PF00481; PP2C; 1.
DR SMART; SM00332; PP2Cc; 1.
DR SMART; SM00331; PP2C-SIG; 1.
DR PROSITE; PS01032; PP2C; 1.
SQ SEQUENCE 377 AA; 41431 MW; 586D46246FB64F5F CRC64;

Query Match      7.5%; Score 193.5; DB 10; Length 377;
Best Local Similarity 24.9%; Pred. No. 1.6e-05;
Matches 77; Conservative 47; Mismatches 114; Indels 71; Gaps 11;

QY 65 YGVFNQYDGNRVNFFVAQRLSAELLQLQNAEADVRVLLQAFDVVERSFLSIDD- 123
Db 111 YGVFDGHGGPEAIAFMKENLT-----RLFFQDA-----VFPMPISIVDAFFLELENS 158
QY 124 -----ALAEKASLSQLPEGVFQHQLPPOYQKILERLKLTEREISGGANAVAVALLNKL 178
Db 159 HRKAFALADLAWADETIVSG-----SCGTTALTALIIGHRL 194
QY 179 YVANVGNTRALLCKSTVDGLQVTLQNVNHTTENEDELFRLSQLG--LDAGKIKQVGIICG 236
Db 195 LVANAGDCRAVLCRRGV-----AVDMSFDRHSYFERRRIEDLGGYFDGYCUNGVLAV-- 248
QY 237 QBSTRIIGYKVKYGYTDIDLSSAAKSPIIAEPIHGAQPLDGVTCFLVLMSEGLYKAL 296
Db 249 ---TRAIGDWELKNPPTD-----SSSLISDPEI-GQILTEDDEFILACDGIWDVL 297
QY 297 EAAHGPQANQ-----ETAMIDTEFAKOTSLDAVAQAVDVRVKRIHSDTFASGGE 347
Db 298 SSQNAVSNVRQGLRRHGDPRQCAMELGKEARLQSSDNMT-VIVICFSSVPSPKQPRR 356
```



```

Db 111 -----OKLIANDRTDLE---SGGSTAVTALLINGKALMIANVDSRAIVSRG 155
QY 195 VDLGQVLTQNVNDHTTENDELFRLSGLDAGKIKQVGIIGQESTRRIIDYKVKYGYTD 254
Db 156 ----KAKQMSVDHDDPDDTERSMIESKG-----GFV-----TNRPGVPRVNGILLA 197
QY 235 IDLLSAASKSP--ITAEPEIHGAQPLDGVGTGLVMSGLVKALEAANGPOANOELIAM 312
Db 198 ISSVGDKKVKAKYLNSEPEIKDV-TIDSHTDFTIIMASDGIKVM-----SNOE---A 245
QY 313 IDTEFAKQTSLDVAQAQVVDVRYKRIHSD 340
Db 246 VDAKKLKDPKEARQVVAEALKNRNSKD 273

RESULT 14
Q9LEW5 PRELIMINARY; PRT; 348 AA.
AC Q9LEW5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Protein phosphatase 2C-like protein.
GN T30N20.10
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxId=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Peters S.A., van Staveren M., Dirkse W., Stiekema W.,
RA Bancroft I., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X.;
RL Submitted (JUL-2000) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Arabidopsis sequencing project;
RL Submitted (JUL-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AL365234; CAB96829.1; -.
DR HSSP; P35813; IAGQ.
DR InterPro; IPR001993; Mitoch_carrier.
DR InterPro; IPR00222; PP2C.
DR InterPro; IPR001932; PP2C-like.
DR Pfam; PF00481; PP2C; 1.
DR SMART; SM00332; PP2Cc; 1.
DR SMART; SM00331; PP2C_SIG; 1.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_1.
DR PROSITE; PS01032; PP2C; 1.
SQ SEQUENCE 348 AA; 37379 MW; 12447BB07F740E2C CRC64;

Query Match 7.4%; Score 190.5; DB 10; Length 348;
Best Local Similarity 20.2%; Pred. No. 2,3e-05;
Matches 94; Conservative 75; Mismatches 155; Indels 141; Gaps 18;

```

```

QY 302 PQANOELIAMT-----DTEFAKQTSLDVAQAQVVDVRYKRIHSDTEFASGERARFCPRHED 357
Db 232 ---SNEAAVAMKEVDEP-----DSAKKLVGEAKIRKGSAD-----N 265
QY 358 MTLVNFPGYPLGEMSQTPSPAPAGRVYPSVPSAOSTSKTSVTLSLWPSOGQM 417
Db 266 ITCVVRF-----LEKSNASSHISSSSSSSEAKEMPEPLGDL 301
QY 418 VNGASHASTID---EATPLINQSPFTLTQSTNTHQSSSSSD 458
Db 302 AISNEAKQVQIGSGNKPENVTNRKPDITASRSTDLTLLERNSTVD 346

RESULT 15
Q9SZN2 PRELIMINARY; PRT; 395 AA.
AC Q9SZN2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical 43.6 kDa protein.
GN P20M13.80 OR A74G38520.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxId=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Medler H., Kutzner M., Wambutt R., Bancroft I., Mewes H.W.,
RA Mayer K.F.X., Schueller C.;
RL Submitted (FEB-1999) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Arabidopsis sequencing project;
RL Submitted (FEB-1999) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Medler H., Kutzner M., Wambutt R., Mewes H.W., Lemcke K.,
RA Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AL035540; CAB37508.1; -.
DR EMBL; AL161593; CAB80516.1; -.
DR InterPro; IPR00222; PP2C.
DR InterPro; IPR001932; PP2C-like.
DR Pfam; PF00481; PP2C; 1.
DR SMART; SM00332; PP2Cc; 1.
DR SMART; SM00331; PP2C_SIG; 1.
DR PROSITE; PS01032; PP2C; 1.
DR Hypothetical protein.
SQ SEQUENCE 395 AA; 43621 MW; 438B4906AFA4C2A CRC64;

Query Match 7.4%; Score 190.5; DB 10; Length 395;
Best Local Similarity 24.5%; Pred. No. 2,8e-05;
Matches 104; Conservative 60; Mismatches 150; Indels 111; Gaps 20;

```

```

Qy 184 GTNRALLCK-STVDG-LQVTQLNVDHHTTENEDELFRLSQLGLDAGKI-----KQVGI 233
Db 163 GDSRAVLGQWVRVTGEAHATQLSAEHNASIESVRRELQALHPDHPDIVVLKHNVRVKGI 222
Qy 234 ICGQESTRIGDYKVKYG-----YTDIDLISAASKPII-ABPEI--HGAQPLDGVGT 283
Db 223 I---QVSRSIGDVYLKRSEFNREPLYAKFRLRSPF-SKPLLSAEPATVHTLEPHD--Q 275
Qy 284 FLVLMSEGLYKALEAAHGFQGOANOEIAAMIDTE----PAKQTSLDAVAQAVVDVKRIHS 339
Db 276 FIICASDGLWEHM-----SNQEAVDIVQNHPRNGIAKRLVKVALQEAACKREMRYSD 327
Qy 340 DTFASGGERARFCPRHEDMTLLVRNFGYPLGEMSQPTSPAPAAAGGRVYPVSVFYSS-AQ 398
Db 328 LKKIDRGVRRHF---HDDITVIVVFFDTNL--VSRGSMRLRGPVSVVRGAGVNLPHNTLAP 382
Qy 399 STSKT 403
Db 383 CTTPT 387

```

Search completed: December 9, 2002, 22:58:26
Job time : 58.4488 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 9, 2002, 22:48:29 ; Search time 21.4072 Seconds
(without alignments)
692.718 Million cell updates/sec

Title: US-09-830-144-4
Perfect score: 2580
Sequence: 1 MAQRRLQLSQPQSWTDD.....AEFVRLMSVDHGEQSWTAP 504

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/1aa/5A-COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B-COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A-COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B-COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCITUS-COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	2580	100.0	504	2	US-08-752-891-2
2	2580	100.0	504	2	US-09-144-178-2
3	2580	100.0	504	2	US-09-406-854-2
4	2580	100.0	504	4	US-09-529-279-2
5	2580	100.0	517	4	US-09-529-279-11
6	2577	99.9	513	4	US-09-529-279-43
7	2575	99.8	504	2	US-08-752-891-6
8	2575	99.8	504	2	US-09-144-178-6
9	2575	99.8	504	4	US-09-406-854-6
10	178	6.9	392	4	US-09-013-881-2
11	125.5	4.9	346	3	US-08-935-855-20
12	124	4.8	306	2	US-08-822-701-8
13	124	4.8	306	3	US-08-935-855-8
14	118	4.6	390	4	US-09-206-646-4
15	118	4.6	392	2	US-08-822-701-2
16	118	4.6	392	3	US-08-935-855-2
17	118	4.6	542	3	US-08-935-855-22
18	117.5	4.6	309	2	US-08-822-701-7
19	117.5	4.6	309	3	US-08-935-855-7
20	117	4.5	390	2	US-08-873-093-3
21	117	4.5	390	4	US-09-206-646-3
22	117	4.5	677	1	US-08-188-582-13
23	117	4.5	677	1	US-08-646-715-13
24	116	4.5	314	2	US-08-822-701-10
25	116	4.5	314	3	US-08-935-855-10
26	110	4.3	479	4	US-09-206-646-1
27	109	4.2	478	2	US-08-873-093-1

28 109 4.2 478 2 US-08-873-093-4 Sequence 4, Appli
29 108 4.2 281 2 US-08-822-701-9 Sequence 9, Appli
30 108 4.2 281 3 US-08-935-855-9 Sequence 9, Appli
31 106 4.1 910 4 US-08-460-269C-2 Sequence 2, Appli
32 105.5 4.1 2930 4 US-09-417-822-2 Sequence 2, Appli
33 105 4.1 1068 4 US-09-085-199B-11 Sequence 11, Appl
34 104.5 4.1 3724 2 US-08-804-227C-10 Sequence 10, Appl
35 104.5 4.1 3724 2 US-08-804-198-4 Sequence 4, Appli
36 103.5 4.0 1066 4 US-09-541-782-8 Sequence 8, Appli
37 103.5 4.0 1066 4 US-09-723-820-8 Sequence 8, Appli
38 103.5 4.0 2471 1 US-08-185-432-16 Sequence 16, Appl
39 103.5 4.0 2471 1 US-08-083-590A-19 Sequence 19, Appl
40 103.5 4.0 2471 3 US-08-532-384-19 Sequence 19, Appl
41 102.5 4.0 2471 4 US-08-899-232-1 Sequence 1, Appli
42 102.5 4.0 594 2 US-08-910-856-2 Sequence 2, Appli
43 100.5 3.9 441 4 US-09-522-666-4 Sequence 4, Appli
44 100.5 3.9 1024 4 US-09-522-666-2 Sequence 2, Appli
45 100.5 3.9 1049 4 US-09-522-666-6 Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-08-752-891-2
; Sequence 2, Application US/08752891
; Patent No. 5837819
; GENERAL INFORMATION:
; APPLICANT: MATSUMOTO, Kunihiro
; APPLICANT: NISHIDA, Eisuke
; TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/752,891
; FILING DATE: 20-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-300856
; FILING DATE: 28-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-126282
; FILING DATE: 24-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 17981/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 504 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-752-891-2

Query Match 100.0%; Score 2580; DB 2; Length 504;
Best Local Similarity 100.0%; Pred. No. 9.6e-230;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MAAQRSLLOSQEQPSWTDLPLCHLSGVGSANRSYSADGKGTSHHPEDSMLKFRSEN 60
D 1 MAAQRSLLOSQEQPSWTDLPLCHLSGVGSANRSYSADGKGTSHHPEDSMLKFRSEN 60
QY 61 NCFLYGVFNQYDGNRVYTNFVAQRLSAELLGQLNAEHAADVRRVLLQAFDVERSFILES 120
D 61 NCFLYGVFNQYDGNRVYTNFVAQRLSAELLGQLNAEHAADVRRVLLQAFDVERSFILES 120
QY 121 IDDALEKASLSQSLPEGVPQHQLPPQYOKILERLKTLEREISGGAMAVAVLLNNKLYV 180
D 121 IDDALEKASLSQSLPEGVPQHQLPPQYOKILERLKTLEREISGGAMAVAVLLNNKLYV 180
QY 181 ANVTGNRALLCSTVDGLQVTLQNVHTTENEDELFRLSQGLDAGKIKOVGIICQEST 240
D 181 ANVTGNRALLCSTVDGLQVTLQNVHTTENEDELFRLSQGLDAGKIKOVGIICQEST 240
QY 241 RRIQDYKVKYGYTDIDLLSAASKPIIAPEIHGAQPLDGVGFLVMSBGLYKALEAAH 300
D 241 RRIQDYKVKYGYTDIDLLSAASKPIIAPEIHGAQPLDGVGFLVMSBGLYKALEAAH 300
QY 301 GPGQANOEIAAMIDTEFAKOTSLDAVAQAVDVVKRIHSDTFASGGERARFCPRHEDMTL 360
D 301 GPGQANOEIAAMIDTEFAKOTSLDAVAQAVDVVKRIHSDTFASGGERARFCPRHEDMTL 360
QY 361 LVNRFQYPLGEMSOPTSPAPAGGRVYPSVYSSAOSTSKTSVTLISLVMSQOGMYNG 420
D 361 LVNRFQYPLGEMSOPTSPAPAGGRVYPSVYSSAOSTSKTSVTLISLVMSQOGMYNG 420
QY 421 AHSASTLDEATPTLTNOSTLTLTOSTNTHTOSSSSSDGGLFRSPAHSLPGEGRVPE 480
D 421 AHSASTLDEATPTLTNOSTLTLTOSTNTHTOSSSSSDGGLFRSPAHSLPGEGRVPE 480
QY 481 YVDFAEFYRLMSVDHGEOQSVVTAP 504
D 481 YVDFAEFYRLMSVDHGEOQSVVTAP 504

```

RESULT 2

US-09-144-178-2
Sequence 2, Application US/09144178

Patent No. 5989862
GENERAL INFORMATION:

APPLICANT: MATSUMOTO, Kunihito

APPLICANT: NISHIDA, Eisuke

TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Foley & Lardner

STREET: 3000 K Street, N.W., Suite 500

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20007-5109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/144,178

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/752,891

FILING DATE: 28-NOV-1996

APPLICATION NUMBER: JP 8-300856

FILING DATE: 28-OCT-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 8-126282

FILING DATE: 24-APR-1996

ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.

```

; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 17981/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; TELEEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 504 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-144-178-2
Query Match 100.0%; Score 2580; DB 2; Length 504;
Best Local Similarity 100.0%; Pred. No. 9.6e-230;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MAAQRSLLOSQEQPSWTDLPLCHLSGVGSANRSYSADGKGTSHHPEDSMLKFRSEN 60
D 1 MAAQRSLLOSQEQPSWTDLPLCHLSGVGSANRSYSADGKGTSHHPEDSMLKFRSEN 60
QY 61 NCFLYGVFNQYDGNRVYTNFVAQRLSAELLGQLNAEHAADVRRVLLQAFDVERSFILES 120
D 61 NCFLYGVFNQYDGNRVYTNFVAQRLSAELLGQLNAEHAADVRRVLLQAFDVERSFILES 120
QY 121 IDDALEKASLSQSLPEGVPQHQLPPQYOKILERLKTLEREISGGAMAVAVLLNNKLYV 180
D 121 IDDALEKASLSQSLPEGVPQHQLPPQYOKILERLKTLEREISGGAMAVAVLLNNKLYV 180
QY 181 ANVTGNRALLCSTVDGLQVTLQNVHTTENEDELFRLSQGLDAGKIKOVGIICQEST 240
D 181 ANVTGNRALLCSTVDGLQVTLQNVHTTENEDELFRLSQGLDAGKIKOVGIICQEST 240
QY 241 RRIQDYKVKYGYTDIDLLSAASKPIIAPEIHGAQPLDGVGFLVMSBGLYKALEAAH 300
D 241 RRIQDYKVKYGYTDIDLLSAASKPIIAPEIHGAQPLDGVGFLVMSBGLYKALEAAH 300
QY 301 GPGQANOEIAAMIDTEFAKOTSLDAVAQAVDVVKRIHSDTFASGGERARFCPRHEDMTL 360
D 301 GPGQANOEIAAMIDTEFAKOTSLDAVAQAVDVVKRIHSDTFASGGERARFCPRHEDMTL 360
QY 361 LVNRFQYPLGEMSOPTSPAPAGGRVYPSVYSSAOSTSKTSVTLISLVMSQOGMYNG 420
D 361 LVNRFQYPLGEMSOPTSPAPAGGRVYPSVYSSAOSTSKTSVTLISLVMSQOGMYNG 420
QY 421 AHSASTLDEATPTLTNOSTLTLTOSTNTHTOSSSSSDGGLFRSPAHSLPGEGRVPE 480
D 421 AHSASTLDEATPTLTNOSTLTLTOSTNTHTOSSSSSDGGLFRSPAHSLPGEGRVPE 480
QY 481 YVDFAEFYRLMSVDHGEOQSVVTAP 504
D 481 YVDFAEFYRLMSVDHGEOQSVVTAP 504

```

RESULT 3

US-09-406-854-2
Sequence 2, Application US/09406854

Patent No. 6140042

GENERAL INFORMATION:

APPLICANT: MATSUMOTO, Kunihito

APPLICANT: NISHIDA, Eisuke

TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Foley & Lardner

STREET: 3000 K Street, N.W., Suite 500

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20007-5109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/406,854
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/752,891
FILING DATE: 20-NOV-1996
APPLICATION NUMBER: JP 8-300856
FILING DATE: 28-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 8-126282
FILING DATE: 24-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 17981/111
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 504 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-406-854-2

Query Match 100.0%; Score 2580; DB 4; Length 504;
Best Local Similarity 100.0%; Pred. No. 9.6e-230;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAAQRSLQSEQPSWTDLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60
Db 1 MAAQRSLQSEQPSWTDLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60
QY 61 NCFLYGVFNQYDGNRVTFNFAQRLSAELLQGLNAEAEADVRVLLQAFDVVERSFLES 120
Db 61 NCFLYGVFNQYDGNRVTFNFAQRLSAELLQGLNAEAEADVRVLLQAFDVVERSFLES 120
QY 121 IDDAEAKASLQSLPEGVPQHQLPPQYQKILRLKTLEREISGGAMAVVAVLLNNKLYV 180
Db 121 IDDAEAKASLQSLPEGVPQHQLPPQYQKILRLKTLEREISGGAMAVVAVLLNNKLYV 180
QY 181 ANVGTNRALLCKSTVDGLQVTLNVDHTTENEDELFRSLQGLDAGKIKQVGIICQEST 240
Db 181 ANVGTNRALLCKSTVDGLQVTLNVDHTTENEDELFRSLQGLDAGKIKQVGIICQEST 240
QY 241 RRIQDYKVKYGYTDIDLLSAKSKPIIAEPEIHGAQPLDGVTFGLVMSEGLYKALEAAH 300
Db 241 RRIQDYKVKYGYTDIDLLSAKSKPIIAEPEIHGAQPLDGVTFGLVMSEGLYKALEAAH 300
QY 301 GPGQANQIEAAMIDTEFAKQTSLDAAVAQAVVDRVKRIHSDTTFASGGERARFCPRHEDMTL 360
Db 301 GPGQANQIEAAMIDTEFAKQTSLDAAVAQAVVDRVKRIHSDTTFASGGERARFCPRHEDMTL 360
QY 361 LVRNFGYPLGEMSOPTSPAPAGGRVYVPVSYSSAQSTSKTSVTLSLVMPSGQMWNG 420
Db 361 LVRNFGYPLGEMSOPTSPAPAGGRVYVPVSYSSAQSTSKTSVTLSLVMPSGQMWNG 420
QY 421 AHSASTLDEATPTLTNOSPPTLTQSTNTHQTSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480
Db 421 AHSASTLDEATPTLTNOSPPTLTQSTNTHQTSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480
QY 481 YVDFAEFYRLWSVDHGQSVVVTAP 504
Db 481 YVDFAEFYRLWSVDHGQSVVVTAP 504

RESULT 4

US-09-529-279-2
Sequence 2, Application US/09529279
Patent No. 6451617
GENERAL INFORMATION:
APPLICANT: ONO, KOICHIRO
APPLICANT: OHTOMO, TOSHIHIKO
APPLICANT: TSUCHIYA, MASAYUKI
TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
FILE REFERENCE: 053466/0278
CURRENT APPLICATION NUMBER: US/09/529,279
CURRENT FILING DATE: 2000-04-11
PRIOR APPLICATION NUMBER: PCT/JP98/04796
PRIOR FILING DATE: 1998-10-22
PRIOR APPLICATION NUMBER: JP 9/290188
PRIOR FILING DATE: 1997-10-22
NUMBER OF SEQ ID NOS: 48
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 2
LENGTH: 504
TYPE: PRT
ORGANISM: Homo sapiens
US-09-529-279-2

Query Match 100.0%; Score 2580; DB 4; Length 504;
Best Local Similarity 100.0%; Pred. No. 9.6e-230;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAAQRSLQSEQPSWTDLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60
Db 1 MAAQRSLQSEQPSWTDLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60
QY 61 NCFLYGVFNQYDGNRVTFNFAQRLSAELLQGLNAEAEADVRVLLQAFDVVERSFLES 120
Db 61 NCFLYGVFNQYDGNRVTFNFAQRLSAELLQGLNAEAEADVRVLLQAFDVVERSFLES 120
QY 121 IDDAEAKASLQSLPEGVPQHQLPPQYQKILRLKTLEREISGGAMAVVAVLLNNKLYV 180
Db 121 IDDAEAKASLQSLPEGVPQHQLPPQYQKILRLKTLEREISGGAMAVVAVLLNNKLYV 180
QY 181 ANVGTNRALLCKSTVDGLQVTLNVDHTTENEDELFRSLQGLDAGKIKQVGIICQEST 240
Db 181 ANVGTNRALLCKSTVDGLQVTLNVDHTTENEDELFRSLQGLDAGKIKQVGIICQEST 240
QY 241 RRIQDYKVKYGYTDIDLLSAKSKPIIAEPEIHGAQPLDGVTFGLVMSEGLYKALEAAH 300
Db 241 RRIQDYKVKYGYTDIDLLSAKSKPIIAEPEIHGAQPLDGVTFGLVMSEGLYKALEAAH 300
QY 301 GPGQANQIEAAMIDTEFAKQTSLDAAVAQAVVDRVKRIHSDTTFASGGERARFCPRHEDMTL 360
Db 301 GPGQANQIEAAMIDTEFAKQTSLDAAVAQAVVDRVKRIHSDTTFASGGERARFCPRHEDMTL 360
QY 361 LVRNFGYPLGEMSOPTSPAPAGGRVYVPVSYSSAQSTSKTSVTLSLVMPSGQMWNG 420
Db 361 LVRNFGYPLGEMSOPTSPAPAGGRVYVPVSYSSAQSTSKTSVTLSLVMPSGQMWNG 420
QY 421 AHSASTLDEATPTLTNOSPPTLTQSTNTHQTSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480
Db 421 AHSASTLDEATPTLTNOSPPTLTQSTNTHQTSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480
QY 481 YVDFAEFYRLWSVDHGQSVVVTAP 504
Db 481 YVDFAEFYRLWSVDHGQSVVVTAP 504

RESULT 5
US-09-529-279-11
Sequence 11, Application US/09529279
Patent No. 6451617
GENERAL INFORMATION:
APPLICANT: ONO, KOICHIRO
APPLICANT: OHTOMO, TOSHIHIKO
APPLICANT: TSUCHIYA, MASAYUKI
TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES

```
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529,279
; CURRENT FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 11
; LENGTH: 517
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-279-11

Query Match      100.0%; Score 2580; DB 4; Length 517;
Best Local Similarity 100.0%; Pred. No. 1e-229;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MAAQRSLLOSQEQPSWTDLPLCHLSGVGSASNSYSADGKGTSHPPEDSWLKFRSEN 60
DB      1 MAAQRSLLOSQEQPSWTDLPLCHLSGVGSASNSYSADGKGTSHPPEDSWLKFRSEN 60
QY      61 NCFLYGVFNGYDGNRVTNFVAQRLSAELLGQLNMAHADVRVLLQAFDVERSFILES 120
DB      61 NCFLYGVFNGYDGNRVTNFVAQRLSAELLGQLNMAHADVRVLLQAFDVERSFILES 120
QY      121 IDDLAEKASLQSLPEGVPOHLPPOYOKILERLKTLEREISGGAMAVAVLLNNKLYV 180
DB      121 IDDLAEKASLQSLPEGVPOHLPPOYOKILERLKTLEREISGGAMAVAVLLNNKLYV 180
QY      121 IDDLAEKASLQSLPEGVPOHLPPOYOKILERLKTLEREISGGAMAVAVLLNNKLYV 180
DB      121 IDDLAEKASLQSLPEGVPOHLPPOYOKILERLKTLEREISGGAMAVAVLLNNKLYV 180
QY      181 ANVTNRALLCKSTVDGLQVTLQNVDTHTENEDLFRLSQLGDAGKIKOVGIIIGQEST 240
DB      181 ANVTNRALLCKSTVDGLQVTLQNVDTHTENEDLFRLSQLGDAGKIKOVGIIIGQEST 240
QY      181 ANVTNRALLCKSTVDGLQVTLQNVDTHTENEDLFRLSQLGDAGKIKOVGIIIGQEST 240
DB      181 ANVTNRALLCKSTVDGLQVTLQNVDTHTENEDLFRLSQLGDAGKIKOVGIIIGQEST 240
QY      241 LVRNFGYPLGMSQPTSPAPAGGRVYVSVYSSAOSTSKTSVTLSLVMSQGMVNG 420
DB      241 LVRNFGYPLGMSQPTSPAPAGGRVYVSVYSSAOSTSKTSVTLSLVMSQGMVNG 420
QY      301 GPGQANOETIAMIDTEFAKOTSLDAVAQAQVDRVKRIHSDTFASGGERARFCPRHEDMTL 360
DB      301 GPGQANOETIAMIDTEFAKOTSLDAVAQAQVDRVKRIHSDTFASGGERARFCPRHEDMTL 360
QY      301 GPGQANOETIAMIDTEFAKOTSLDAVAQAQVDRVKRIHSDTFASGGERARFCPRHEDMTL 360
DB      301 GPGQANOETIAMIDTEFAKOTSLDAVAQAQVDRVKRIHSDTFASGGERARFCPRHEDMTL 360
QY      361 LVRNFGYPLGMSQPTSPAPAGGRVYVSVYSSAOSTSKTSVTLSLVMSQGMVNG 420
DB      361 LVRNFGYPLGMSQPTSPAPAGGRVYVSVYSSAOSTSKTSVTLSLVMSQGMVNG 420
QY      421 AHSASTLDEATPTLTNOSPTLTLOSTNTHTOSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480
DB      421 AHSASTLDEATPTLTNOSPTLTLOSTNTHTOSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480
QY      481 YVDFAEFYRLMSVDHGQSVVTAP 504
DB      481 YVDFAEFYRLMSVDHGQSVVTAP 504

RESULT 6
US-09-529-279-43
; Sequence 43, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529,279
; CURRENT FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
```

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 43
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-279-43

Query Match      99.8%; Score 2577; DB 4; Length 513;
Best Local Similarity 99.8%; Pred. No. 1.9e-229;
Matches 503; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 MAAQRSLLOSQEQPSWTDLPLCHLSGVGSASNSYSADGKGTSHPPEDSWLKFRSEN 60
DB      10 MAAQRSLLOSQEQPSWTDLPLCHLSGVGSASNSYSADGKGTSHPPEDSWLKFRSEN 69
QY      61 NCFLYGVFNGYDGNRVTNFVAQRLSAELLGQLNMAHADVRVLLQAFDVERSFILES 120
DB      70 NCFLYGVFNGYDGNRVTNFVAQRLSAELLGQLNMAHADVRVLLQAFDVERSFILES 129
QY      121 IDDLAEKASLQSLPEGVPOHLPPOYOKILERLKTLEREISGGAMAVAVLLNNKLYV 180
DB      130 IDDLAEKASLQSLPEGVPOHLPPOYOKILERLKTLEREISGGAMAVAVLLNNKLYV 189
QY      181 ANVTNRALLCKSTVDGLQVTLQNVDTHTENEDLFRLSQLGDAGKIKOVGIIIGQEST 240
DB      190 ANVTNRALLCKSTVDGLQVTLQNVDTHTENEDLFRLSQLGDAGKIKOVGIIIGQEST 249
QY      241 LVRNFGYPLGMSQPTSPAPAGGRVYVSVYSSAOSTSKTSVTLSLVMSQGMVNG 420
DB      250 LVRNFGYPLGMSQPTSPAPAGGRVYVSVYSSAOSTSKTSVTLSLVMSQGMVNG 429
QY      301 GPGQANOETIAMIDTEFAKOTSLDAVAQAQVDRVKRIHSDTFASGGERARFCPRHEDMTL 360
DB      310 GPGQANOETIAMIDTEFAKOTSLDAVAQAQVDRVKRIHSDTFASGGERARFCPRHEDMTL 369
QY      361 LVRNFGYPLGMSQPTSPAPAGGRVYVSVYSSAOSTSKTSVTLSLVMSQGMVNG 420
DB      370 LVRNFGYPLGMSQPTSPAPAGGRVYVSVYSSAOSTSKTSVTLSLVMSQGMVNG 429
QY      421 AHSASTLDEATPTLTNOSPTLTLOSTNTHTOSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480
DB      430 AHSASTLDEATPTLTNOSPTLTLOSTNTHTOSSSSSDGGLFRSRPAHSLPPGEDGRVEP 489
QY      481 YVDFAEFYRLMSVDHGQSVVTAP 504
DB      490 YVDFAEFYRLMSVDHGQSVVTAP 513

RESULT 7
US-08-752-891-6
; Sequence 6, Application US/08752891
; Patent No. 5837819
; GENERAL INFORMATION:
; APPLICANT: MATSUMOTO, Kunihiko
; APPLICANT: NISHIDA, Eisuke
; TITLE OF INVENTION: TABI PROTEIN AND DNA CODING THEREFOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/752,891
; FILING DATE: 20-NOV-1996
; CLASSIFICATION: 435
```

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-300856
; FILING DATE: 28-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-126282
; FILING DATE: 24-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 17981/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 504 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-752-891-6

Query Match          99.8%; Score 2575; DB 2; Length 504;
Best Local Similarity 99.8%; Pred. No. 2.8e-229;
Matches 503; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAAQRSLLOEQSPSWTDDLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60
DB 1 MAAQRSLLOEQSPSWTDDLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60
QY 61 NCFLYGVFNGYDGNRVNFTNFAQRLSAELLQLGNAEHAEDVRRVLLQAPDVVERSFLES 120
DB 61 NCFLYGVFNGYDGNRVNFTNFAQRLSAELLQLGNAEHAEDVRRVLLQAPDVVERSFLES 120
QY 121 IDDALEAKASLOSQLPBGVPOHQLPPQYQKILERLKTLEIREISGGAMAVVAVLLNNKLYV 180
DB 121 IDDALEAKASLOSQLPBGVPOHQLPPQYQKILERLKTLEIREISGGAMAVVAVLLNNKLYV 180
QY 181 ANVGTNRALLCKSTVDGLQVTLQNVNHTTENEDELFRLSQLGLDAGKIKQVGIICQEST 240
DB 181 ANVGTNRALLCKSTVDGLQVTLQNVNHTTENEDELFRLSQLGLDAGKIKQVGIICQEST 240
QY 241 RRTGDKYVKYGYTDIDLLSAASKPIIAEPEIHGAQPLDGVTFGLVMSGLYKALEAAH 300
DB 241 RRTGDKYVKYGYTDIDLLSAASKPIIAEPEIHGAQPLDGVTFGLVMSGLYKALEAAH 300
QY 301 GPGQANQEIAMIDTEFAKOTSLDAVAQAVVDRVKRIHSDTFASGGERARFCPRHEDMTL 360
DB 301 GPGQANQEIAMIDTEFAKOTSLDAVAQAVVDRVKRIHSDTFASGGERARFCPRHEDMTL 360
QY 361 LVRNFGYPLGEMSOPTSPAPAGGRVYVPVYSSAQSTSKTSVTLSLVMPSGQMVNG 420
DB 361 LVRNFGYPLGEMSOPTSPAPAGGRVYVPVYSSAQSTSKTSVTLSLVMPSGQMVNG 420
QY 421 AHSASTLDEATPTLTNQSPTLTLOSTNTHQTQSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480
DB 421 AHSASTLDEATPTLTNQSPTLTLOSTNTHQTQSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480
QY 481 YVDFAEFYRLWSVDHGQSVVVTAP 504
DB 481 YVDFAEFYRLWSVDHGQSVVVTAP 504

RESULT 8
US-09-144-178-6
; Sequence 6, Application US/09144178
; Patent No. 5989862
; GENERAL INFORMATION:
; APPLICANT: MATSUMOTO, Kunihiro
; APPLICANT: NISHIDA, Eisuke
; TITLE OF INVENTION: TABI PROTEIN AND DNA CODING THEREFOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
```

```
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/144.178
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/752,891
; FILING DATE: 20-NOV-1996
; APPLICATION NUMBER: JP 8-300856
; FILING DATE: 28-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-126282
; FILING DATE: 24-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 17981/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 504 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-144-178-6

Query Match          99.8%; Score 2575; DB 2; Length 504;
Best Local Similarity 99.8%; Pred. No. 2.8e-229;
Matches 503; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAAQRSLLOEQSPSWTDDLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60
DB 1 MAAQRSLLOEQSPSWTDDLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60
QY 61 NCFLYGVFNGYDGNRVNFTNFAQRLSAELLQLGNAEHAEDVRRVLLQAPDVVERSFLES 120
DB 61 NCFLYGVFNGYDGNRVNFTNFAQRLSAELLQLGNAEHAEDVRRVLLQAPDVVERSFLES 120
QY 121 IDDALEAKASLOSQLPBGVPOHQLPPQYQKILERLKTLEIREISGGAMAVVAVLLNNKLYV 180
DB 121 IDDALEAKASLOSQLPBGVPOHQLPPQYQKILERLKTLEIREISGGAMAVVAVLLNNKLYV 180
QY 181 ANVGTNRALLCKSTVDGLQVTLQNVNHTTENEDELFRLSQLGLDAGKIKQVGIICQEST 240
DB 181 ANVGTNRALLCKSTVDGLQVTLQNVNHTTENEDELFRLSQLGLDAGKIKQVGIICQEST 240
QY 241 RRTGDKYVKYGYTDIDLLSAASKPIIAEPEIHGAQPLDGVTFGLVMSGLYKALEAAH 300
DB 241 RRTGDKYVKYGYTDIDLLSAASKPIIAEPEIHGAQPLDGVTFGLVMSGLYKALEAAH 300
QY 301 GPGQANQEIAMIDTEFAKOTSLDAVAQAVVDRVKRIHSDTFASGGERARFCPRHEDMTL 360
DB 301 GPGQANQEIAMIDTEFAKOTSLDAVAQAVVDRVKRIHSDTFASGGERARFCPRHEDMTL 360
QY 361 LVRNFGYPLGEMSOPTSPAPAGGRVYVPVYSSAQSTSKTSVTLSLVMPSGQMVNG 420
DB 361 LVRNFGYPLGEMSOPTSPAPAGGRVYVPVYSSAQSTSKTSVTLSLVMPSGQMVNG 420
QY 421 AHSASTLDEATPTLTNQSPTLTLOSTNTHQTQSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480
DB 421 AHSASTLDEATPTLTNQSPTLTLOSTNTHQTQSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480
```

Db 421 AHSASTLDEATPTLTNQSPTLTLTOSTNTHTOSSSSSDGGLFRSRPAHSLPPGEDGRVER 480
QY 481 YVDFAEFYRLMSVDHGQSVVTAP 504
Db 481 YVDFAEFYRLMSVDHGQSVVTAP 504

RESULT 9
US-09-406-854-6

; Sequence 6, Application US/09406854
; Patent No. 6140042
; GENERAL INFORMATION:
; APPLICANT: MATSUMOTO, Kunihiko
; APPLICANT: NISHIDA, Eisuke
; TITLE OF INVENTION: TABI PROTEIN AND DNA CODING THEREFOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/406,854
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/752,891
; FILING DATE: 20-NOV-1996
; APPLICATION NUMBER: JP 8-300856
; FILING DATE: 28-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-126282
; FILING DATE: 24-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 17981/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 504 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-406-854-6

Query Match 99.8%; Score 2575; DB 4; Length 504;
Best Local Similarity 99.8%; Pred. No. 2,8e-229;
Matches 503; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAQRSLLOSSEQPSTWTDLPICHSIGVSASNRYSADGKTBESHPPEDSWLKRSEN 60
Db 1 MAQRSLLOSSEQPSTWTDLPICHSIGVSASNRYSADGKTBESHPPEDSWLKRSEN 60
QY 61 NCFYGVFNGYDGNRYTNFPAORLSAELLGOLNAEHAADYRVYLAQFDVVERSFLES 120
Db 61 NCFYGVFNGYDGNRYTNFPAORLSAELLGOLNAEHAADYRVYLAQFDVVERSFLES 120
QY 121 IDBALAEKASLOSQLEPGVPOHOLPPYOYKILERLKTLEREISGMAVAVALNKKLYV 180
Db 121 IDBALAEKASLOSQLEPGVPOHOLPPYOYKILERLKTLEREISGMAVAVALNKKLYV 180
QY 181 ANVTNRALLCKSTVVDGLOVTLNVDDHTTENEDLFRLSQGLDAGKIKQVIGIGQEST 240

Db 181 ANVTNRALLCKSTVVDGLOVTLNVDDHTTENEDLFRLSQGLDAGKIKQVIGIGQEST 240
QY 241 RRTGDKYKVGKGTIDIDLSAAKSKPIAEPETHGQPLDGVTFVLVSEGLYKALEAAH 300
Db 241 RRTGDKYKVGKGTIDIDLSAAKSKPIAEPETHGQPLDGVTFVLVSEGLYKALEAAH 300
QY 301 GPGQANQEIAMIDTEFAKQTSIDAVAQAVVDVRVRIHSDTFASGGERARFCPRHEDMTL 360
Db 301 GPGQANQEIAMIDTEFAKQTSIDAVAQAVVDVRVRIHSDTFASGGERARFCPRHEDMTL 360
QY 361 LVNFGYPIGEMSQPTSPAPAAAGRVYVSVYPSAOSTKTSVTLSLVMSQGMWNG 420
Db 361 LVNFGYPIGEMSQPTSPAPAAAGRVYVSVYPSAOSTKTSVTLSLVMSQGMWNG 420
QY 421 AHSASTLDEATPTLTNQSPTLTLTOSTNTHTOSSSSSDGGLFRSRPAHSLPPGEDGRVER 480
Db 421 AHSASTLDEATPTLTNQSPTLTLTOSTNTHTOSSSSSDGGLFRSRPAHSLPPGEDGRVER 480
QY 481 YVDFAEFYRLMSVDHGQSVVTAP 504
Db 481 YVDFAEFYRLMSVDHGQSVVTAP 504

RESULT 10

US-09-013-881-2
; Sequence 2, Application US/09013881
; Patent No. 6132964
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; TITLE OF INVENTION: HUMAN HYDROLASE-LIKE MOLECULES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/013,881
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: BILLINGS, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0470 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 392 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: KIDNOT02
; CLONE: 195647

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/822,701
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 1049-1-002 N
TELECOMMUNICATION INFORMATION:

```

; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 306 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM: Rattus
;
US-08-822-701-8

Query Match          4.8%; Score 124; DB 2; Length 306;
Best Local Similarity 19.3%; Pred. No. 0.0029;
Matches 53; Conservative 49; Mismatches 110; Indels 62; Gaps 8;
```

```

QY 65 YGVFNGYDGNRYTNFVAQRLSAELLGQLNAEHAADVRRVLLQAFDVERSFLSIDA 124
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 56 FAVYDGHAGSQVAKYCCCHL-----LDHITNN 82
QY 125 LAEKASLSQQLPEGVPOHLPPOYOKILERLKTLEIREISG---GAMAVVAVLNNKLY 180
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 83 QDFKSGAGAPSVENV-KNGIRTGFLIEDHMRVMSKKHGAADRSSGTAVGLISPOHTYF 141
QY 181 ANVTNRALLCKSTVDGLQVTLNVHTTENEDLFRLSQLGLDAGKIKOVGIICQEST 240
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 142 INGDSSRGLLCNR---KVHFTQDHKPSNPLEKERIQNAG---GSVMIGRVNSGLAVS 194
QY 241 RRIG--DYKVKYGYTDIDLLSAKSKPITAEPEIHGAQPLDGTGFLVMSEGLYKALEA 298
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 195 RALGDPDYKCVHGKGPTEQL-----VSPEPEVHDIERSEDDQFIILACGIDWVM-- 245
QY 299 AHGPGQANOETIAMIDTEFAKQTSLDVAVAQAVVD 332
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 246 -----GNEELCDPVRSRLVETDLEKVCNEVVD 273
```

```

RESULT 13
US-08-935-855-8
; Sequence 8, Application US/08935855
; Patent No. 6066485
; GENERAL INFORMATION:
; APPLICANT: Guttridge, Mark
; APPLICANT: Basilio, Claudio
; TITLE OF INVENTION: NOVEL GROWTH FACTOR INDUCIBLE
; TITLE OF INVENTION: SERINE/THREONINE PHOSPHATASE, FIN13
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/935,855
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1049-1-002 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
```

```

; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 306 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM: Rattus
;
US-08-935-855-8

Query Match          4.8%; Score 124; DB 3; Length 306;
Best Local Similarity 19.3%; Pred. No. 0.0029;
Matches 53; Conservative 49; Mismatches 110; Indels 62; Gaps 8;
```

```

QY 65 YGVFNGYDGNRYTNFVAQRLSAELLGQLNAEHAADVRRVLLQAFDVERSFLSIDA 124
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 56 FAVYDGHAGSQVAKYCCCHL-----LDHITNN 82
QY 125 LAEKASLSQQLPEGVPOHLPPOYOKILERLKTLEIREISG---GAMAVVAVLNNKLY 180
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 83 QDFKSGAGAPSVENV-KNGIRTGFLIEDHMRVMSKKHGAADRSSGTAVGLISPOHTYF 141
QY 181 ANVTNRALLCKSTVDGLQVTLNVHTTENEDLFRLSQLGLDAGKIKOVGIICQEST 240
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 142 INGDSSRGLLCNR---KVHFTQDHKPSNPLEKERIQNAG---GSVMIGRVNSGLAVS 194
QY 241 RRIG--DYKVKYGYTDIDLLSAKSKPITAEPEIHGAQPLDGTGFLVMSEGLYKALEA 298
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 195 RALGDPDYKCVHGKGPTEQL-----VSPEPEVHDIERSEDDQFIILACGIDWVM-- 245
QY 299 AHGPGQANOETIAMIDTEFAKQTSLDVAVAQAVVD 332
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 246 -----GNEELCDPVRSRLVETDLEKVCNEVVD 273
```

```

RESULT 14
US-09-206-646-4
; Sequence 4, Application US/09206646
; Patent No. 6436637
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; APPLICANT: Lal, Preeti G.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: NEW PROTEIN PHOSPHATASE
; FILE REFERENCE: PF-0319-1 DIV
; CURRENT APPLICATION NUMBER: US/09/206,646
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ. ID NOS: 4
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: ID No. 6436637 g452526
;
US-09-206-646-4
```

```

Query Match          4.6%; Score 118; DB 4; Length 390;
Best Local Similarity 21.4%; Pred. No. 0.016;
Matches 60; Conservative 52; Mismatches 107; Indels 62; Gaps 12;

QY 60 NNCFYGVFNGYDGNRYTNFVAQRLSAELLGQLNAEHAADVRRVLLQAFDVERSFLS 119
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 51 DNWSFPAVYDGHAGSRVANCSTHLEHI---TTNEDFRAAD-----KSGSLABS-VE 100
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 120 SIDDALAEKASLSQQLPEGVPOHLPPOYOKILERLKTLEIREISG---GAMAVVAVLNN 175
```

[illegible]

```

RESULT 15
US-08-822-701-2
; Sequence 2, Application US/08822701
; Patent No. 5976853
; GENERAL INFORMATION:
; APPLICANT: Guthridge, Mark
; APPLICANT: Basilico, Claudio
; TITLE OF INVENTION: NOVEL GROWTH FACTOR INDUCIBLE
; TITLE OF INVENTION: SERINE/THREONINE PHOSPHATASE, FINI3
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; STREET: Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/822,701
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1049-1-002 N
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 392 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FRAGMENT TYPE: N-terminal
US-08-822-701-2

```

Query Match 4.6%; Score 118; DB 2; Length 392;
Best Local Similarity 22.2%; Pred. No. 0.016;
Matches 43; Conservative 44; Mismatches 65; Indels 42; Gaps 8;

[illegible]

```
QY 278 LDGVTVFLVMSGLYKALEAHGPGGANOEIAAMTDTFAKOTS-----LDAVAQVV 331
      :|||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 276 LTDDHEFWJACDGNVM-----SSQEVVDFFIQSKIQRDENGELRLUSSIVEELL 327
      :|||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
QY 332 DRVKRIHSDTFASG 345
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 328 DQC--LAPDTSGDG 339
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
```

Search completed: December 9, 2002, 23:00:38
Job time : 23.4072 secs

Search completed: December 9, 2002, 23:00:38
Job time : 23.4072 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 9, 2002, 22:53:24 ; Search time 100.986 Seconds
(without alignments)
81.062 Million cell updates/sec

Title: US-09-830-144-4
Perfect score: 2580
Sequence: 1 MAAQRSLQSEQQPSWTD.....AEFYRLWSVDHGQSYVTAP 504

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 103943 seqs, 16242309 residues

Total number of hits satisfying chosen parameters: 103943

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US05_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2580	100.0	504	9	US-10-158-895-2
2	2580	100.0	504	12	US-10-123-427-2
3	2580	100.0	517	9	US-10-158-895-11
4	2577	99.9	513	9	US-10-158-895-43
5	2575	99.8	504	12	US-10-123-427-6
6	409	15.9	84	10	US-09-923-300-1270
7	365	14.1	70	10	US-09-864-761-34065
8	261	10.1	51	10	US-09-864-761-34067
9	205	7.9	274	10	US-09-860-351-4
10	178	6.9	392	9	US-09-915-124A-2
11	170.5	6.6	372	9	US-09-973-941-4
12	170.5	6.6	372	10	US-09-973-963-4
13	170.5	6.6	372	10	US-09-973-064-4
14	170.5	6.6	372	10	US-09-973-077-4
15	170.5	6.6	372	10	US-09-973-063-4
16	170.5	6.6	372	10	US-09-973-964-4
17	170.5	6.6	372	10	US-09-975-072-4
18	170.5	6.6	372	10	US-09-972-038-4
19	170.5	6.6	372	10	US-09-972-757-4

20	170.5	6.6	372	10	US-09-973-965-4	Sequence 4, Appli
21	164.5	6.4	352	10	US-09-860-351-2	Sequence 2, Appli
22	164.5	6.4	373	10	US-09-925-300-1655	Sequence 1655, Ap
23	156.5	6.1	353	10	US-09-828-302-14	Sequence 14, Appl
24	141.5	5.5	371	10	US-09-828-302-15	Sequence 15, Appl
25	132	5.1	228	10	US-09-900-715-2	Sequence 2, Appli
26	125.5	4.9	387	10	US-09-866-987-9	Sequence 9, Appli
27	125	4.8	382	10	US-09-972-741-2	Sequence 2, Appli
28	125	4.8	1080	10	US-09-781-558-4	Sequence 4, Appli
29	125	4.8	1115	9	US-10-106-534-7	Sequence 7, Appli
30	125	4.8	1135	10	US-09-737-149-35	Sequence 35, Appl
31	121.5	4.7	149	10	US-09-860-351-6	Sequence 6, Appli
32	121	4.7	1031	10	US-09-815-242-10932	Sequence 10932, A
33	117.5	4.6	459	10	US-09-866-987-8	Sequence 8, Appli
34	117	4.5	390	9	US-10-072-130-3	Sequence 3, Appli
35	116	4.5	1115	10	US-09-781-558-2	Sequence 2, Appli
36	114.5	4.4	798	9	US-09-978-295A-405	Sequence 405, App
37	114.5	4.4	798	9	US-09-978-697-405	Sequence 405, App
38	114.5	4.4	798	9	US-09-978-192A-405	Sequence 405, App
39	114.5	4.4	1289	9	US-09-932-145-11	Sequence 11, Appl
40	110.5	4.3	400	10	US-09-925-300-1254	Sequence 1254, Ap
41	109	4.2	478	9	US-10-072-130-1	Sequence 1, Appli
42	109	4.2	478	9	US-10-072-130-4	Sequence 4, Appli
43	108	4.2	281	10	US-09-801-368-278	Sequence 278, App
44	103.5	4.0	1230	10	US-09-881-752A-150	Sequence 150, App
45	102	4.0	143	10	US-09-866-987-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-10-158-895-2
; Sequence 2, Application US/10158895
; Patent No. US20020155624A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHITAKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-158-895-2

Query Match 100.0%; Score 2580; DB 9; Length 504;
Best Local Similarity 100.0%; Pred. No. 5e-194;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAAQRSLQSEQQPSWTD	DLPLCHLSCVGSASNRYSADGKGTESHPPEDSWLKFRSEN	60
Db	1	MAAQRSLQSEQQPSWTD	DLPLCHLSCVGSASNRYSADGKGTESHPPEDSWLKFRSEN	60
QY	61	NCFLYGVFNVDGNRVTFV	QAQLSAELLQLQNAEADRVRLVLLQAFDVSFLES	120
Db	61	NCFLYGVFNVDGNRVTFV	QAQLSAELLQLQNAEADRVRLVLLQAFDVSFLES	120
QY	121	IDDALAEKASIQSQLP	EGVPHQPPQYQKILRLKTLERISGAMAVAVLLNNKLYV	180
Db	121	IDDALAEKASIQSQLP	EGVPHQPPQYQKILRLKTLERISGAMAVAVLLNNKLYV	180

QY 181 ANVTNRALLCKSTVDGLQVTLQNVNHTTENEDELFRLSQGLDAGKIKQVGIICQEST 240
 DB 181 ANVTNRALLCKSTVDGLQVTLQNVNHTTENEDELFRLSQGLDAGKIKQVGIICQEST 240
 QY 241 RRIQDYKVKYGTIDDLISAASKPIIAEPEIHGAOPLDGVTFVLMSEGLYKALEAAH 300
 DB 241 RRIQDYKVKYGTIDDLISAASKPIIAEPEIHGAOPLDGVTFVLMSEGLYKALEAAH 300
 QY 301 GGGQANOEIAAMIDTEFAKQTSIDAVAQAQVVDVKRIHSDTFASGGERARFCPRHEDMTL 360
 DB 301 GGGQANOEIAAMIDTEFAKQTSIDAVAQAQVVDVKRIHSDTFASGGERARFCPRHEDMTL 360
 QY 361 LVNRFQYPLGEMSOPTPSPAPAAAGRVYPSVYSSAOSTSKTSVTLSLVMPSQGMVNG 420
 DB 361 LVNRFQYPLGEMSOPTPSPAPAAAGRVYPSVYSSAOSTSKTSVTLSLVMPSQGMVNG 420
 QY 421 AHSASTLDEATPTLTNOSTPTLTQSTNTHHTOSSSSSDGGLFRSPAHSLPGEEDRVER 480
 DB 421 AHSASTLDEATPTLTNOSTPTLTQSTNTHHTOSSSSSDGGLFRSPAHSLPGEEDRVER 480
 QY 481 YVDFAEFYRLMSVDHGEGSVVTAP 504
 DB 481 YVDFAEFYRLMSVDHGEGSVVTAP 504

RESULT 2

US-10-123-427-2
 ; Sequence 2, Application US/10123427
 ; Patent No. US20020119525A1

GENERAL INFORMATION:

APPLICANT: MATSUMOTO, Kunihito

NISHIDA, Eisuke

TITLE OF INVENTION: TABI PROTEIN AND DNA CODING THEREFOR

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Foley & Lardner

STREET: 3000 K Street, N.W., Suite 500

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20007-5109

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/123,427

FILING DATE: 17-Apr-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/406,854

FILING DATE: <Unknown>

APPLICATION NUMBER: US/08/752,891

FILING DATE: 20-NOV-1996

APPLICATION NUMBER: JP 8-300856

FILING DATE: 28-OCT-1996

APPLICATION NUMBER: JP 8-126282

FILING DATE: 24-APR-1996

ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.

REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 17981/111

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)672-5300

TELEFAX: (202)672-5399

TELEX: 904136

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 504 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 US-10-123-427-2

Query Match 100.0%; Score 2580; DB 12; Length 504;

Best Local Similarity 100.0%; Pred. No. 5e-194;

Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAORSLSLOEQPSWTDLPLCHLSVGSASNSRYSADGKGTESHPPEDSWLKFRSEN 60
 DB 1 MAAORSLSLOEQPSWTDLPLCHLSVGSASNSRYSADGKGTESHPPEDSWLKFRSEN 60
 QY 61 NCFLYGVFNQYDGNRYTNFAORLSAELLGOLNAEHAADYRVLLOAFDVERSFILES 120
 DB 61 NCFLYGVFNQYDGNRYTNFAORLSAELLGOLNAEHAADYRVLLOAFDVERSFILES 120
 QY 121 IDDALEKASLQSLPEGVPOHLPPOYOKILERLKTLEREISGAMAAVAVLNNKLYV 180
 DB 121 IDDALEKASLQSLPEGVPOHLPPOYOKILERLKTLEREISGAMAAVAVLNNKLYV 180
 QY 181 ANVTNRALLCKSTVDGLQVTLQNVNHTTENEDELFRLSQGLDAGKIKQVGIICQEST 240
 DB 181 ANVTNRALLCKSTVDGLQVTLQNVNHTTENEDELFRLSQGLDAGKIKQVGIICQEST 240
 QY 241 RRIQDYKVKYGTIDDLISAASKPIIAEPEIHGAOPLDGVTFVLMSEGLYKALEAAH 300
 DB 241 RRIQDYKVKYGTIDDLISAASKPIIAEPEIHGAOPLDGVTFVLMSEGLYKALEAAH 300
 QY 301 GGGQANOEIAAMIDTEFAKQTSIDAVAQAQVVDVKRIHSDTFASGGERARFCPRHEDMTL 360
 DB 301 GGGQANOEIAAMIDTEFAKQTSIDAVAQAQVVDVKRIHSDTFASGGERARFCPRHEDMTL 360
 QY 361 LVNRFQYPLGEMSOPTPSPAPAAAGRVYPSVYSSAOSTSKTSVTLSLVMPSQGMVNG 420
 DB 361 LVNRFQYPLGEMSOPTPSPAPAAAGRVYPSVYSSAOSTSKTSVTLSLVMPSQGMVNG 420
 QY 421 AHSASTLDEATPTLTNOSTPTLTQSTNTHHTOSSSSSDGGLFRSPAHSLPGEEDRVER 480
 DB 421 AHSASTLDEATPTLTNOSTPTLTQSTNTHHTOSSSSSDGGLFRSPAHSLPGEEDRVER 480
 QY 481 YVDFAEFYRLMSVDHGEGSVVTAP 504
 DB 481 YVDFAEFYRLMSVDHGEGSVVTAP 504

RESULT 3

US-10-158-895-11
 ; Sequence 11, Application US/10158895
 ; Patent No. US20020155624A1

GENERAL INFORMATION:

APPLICANT: ONO, KOICHIRO

APPLICANT: OHTOMO, TOSHIHIKO

APPLICANT: TSUCHIYA, MASAYUKI

TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES

FILE REFERENCE: 053466/0278

CURRENT APPLICATION NUMBER: US/10/158,895

PRIOR FILING DATE: 2002-06-03

PRIOR APPLICATION NUMBER: US/09/529,279

PRIOR FILING DATE: 2000-04-11

PRIOR APPLICATION NUMBER: PCT/JP98/04796

PRIOR FILING DATE: 1998-10-22

PRIOR APPLICATION NUMBER: JP 9/290188

PRIOR FILING DATE: 1997-10-22

NUMBER OF SEQ ID NOS: 48

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 11

LENGTH: 517

TYPE: PRT

ORGANISM: Homo sapiens

US-10-158-895-11

Query Match 100.0%; Score 2580; DB 9; Length 517;
 Best Local Similarity 100.0%; Pred. No. 5.2e-194;
 Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MAAQRRLLOSQEQP	SWTDDLP	CHLSGVGSGASNR	SYADGKTEGHP	PEDSMLKPR	SRN 60
Db	1	MAAQRRLLOSQEQP	SWTDDLP	CHLSGVGSGASNR	SYADGKTEGHP	PEDSMLKPR	SRN 60
Qy	61	NCFLYGVFN	YDGNRTNF	VAQRLSAE	LLGQLNAE	HAEDVRRV	LLQAFDVVERSFLES 120
Db	61	NCFLYGVFN	YDGNRTNF	VAQRLSAE	LLGQLNAE	HAEDVRRV	LLQAFDVVERSFLES 120
Qy	121	IDDALAEKAS	LSQOLPEGV	POHOLP	POYQKIL	ERLKT	LEREISGGMAVVVLLNNKLXV 180
Db	121	IDDALAEKAS	LSQOLPEGV	POHOLP	POYQKIL	ERLKT	LEREISGGMAVVVLLNNKLXV 180
Qy	181	ANVGTNRALL	CKSTVDGL	QVLTQ	LVNDHT	TENED	FLRSLQLGLDAGKIKQVGIIICQEST 240
Db	181	ANVGTNRALL	CKSTVDGL	QVLTQ	LVNDHT	TENED	FLRSLQLGLDAGKIKQVGIIICQEST 240
Qy	241	RRIGDYKV	KVGYTDD	ILL	SAASKP	IIAE	PEIIGAOPDGVTFGLVMSEGLYKALEAAH 300
Db	241	RRIGDYKV	KVGYTDD	ILL	SAASKP	IIAE	PEIIGAOPDGVTFGLVMSEGLYKALEAAH 300
Qy	301	GPGQANQEI	AAMIDTE	FAKOTS	LDVAQA	VVDVR	KRIHSDTFASGGERARFCPRHEDMTL 360
Db	301	GPGQANQEI	AAMIDTE	FAKOTS	LDVAQA	VVDVR	KRIHSDTFASGGERARFCPRHEDMTL 360
Qy	361	LVRNFGY	PLGEMSO	PTSPAPA	AGGRVY	PVSP	YSSAQSTKTSVTLSLWPSQOGWNG 420
Db	361	LVRNFGY	PLGEMSO	PTSPAPA	AGGRVY	PVSP	YSSAQSTKTSVTLSLWPSQOGWNG 420
Qy	421	AHSASTL	DEATPTLT	NOSPTLT	LQSTNT	HTQSSSS	SDGLFRSRPAHSLPPGEDGRVEP 480
Db	421	AHSASTL	DEATPTLT	NOSPTLT	LQSTNT	HTQSSSS	SDGLFRSRPAHSLPPGEDGRVEP 480
Qy	481	YVDFAE	FYRLWS	VDHGEQ	SVV	TAP 504	
Db	481	YVDFAE	FYRLWS	VDHGEQ	SVV	TAP 504	

```

RESULT 4
US-10-158-895-43
; Sequence 43, Application US/10158895
; Patent No. US20020155624A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-158-895-43

```

Db	70	NCFLYGVNGYDGNRVTFNVAQRLSAEILLQGLNABHADVRRVLLQAPDVVHRSFLS	120
Qy	121	IDDALEAKASIQSLPEGVPOHQLPPQYQKILERLKTLEREISGGGAMAVVAVLLNNKLYV	180
Db	130	IDDALEAKASIQSLPEGVPOHQLPPQYQKILERLKTLEREISGGGAMAVVAVLLNNKLYV	189
Qy	181	ANVGTNRALLCKSTVGLQVTLQNVNVDHTTENEDELFRLSQLGLDAGKI KOVGIIICQEST	240
Db	190	ANVGTNRALLCKSTVGLQVTLQNVNVDHTTENEDELFRLSQLGLDAGKI KOVGIIICQEST	249
Qy	241	RRIGDYKVKYGYTDIDLLSAASKPPIIABPEIHGAQPLDQVGTGLVLMSEGLYKALEAAH	300
Db	250	RRIGDYKVKYGYTDIDLLSAASKPPIIABPEIHGAQPLDQVGTGLVLMSEGLYKALEAAH	309
Qy	301	GPQANQOEIAAMIDTEFAKOTSLDVAQAQVVDVRKRIHSDTTFASGGERARFCPRHEDMTL	360
Db	310	GPQANQOEIAAMIDTEFAKOTSLDVAQAQVVDVRKRIHSDTTFASGGERARFCPRHEDMTL	369
Qy	361	LVRNFGYPLGEMSQPTSPAPAGGRVYPVSPYSSAQSTKSTVTLSLVMPSQGQWNG	420
Db	370	LVRNFGYPLGEMSQPTSPAPAGGRVYPVSPYSSAQSTKSTVTLSLVMPSQGQWNG	429
Qy	421	AHSASTLDEATPLTWNQSPFLTILQSTNTHHTQSSSSSSDGLFRSPRPAHSLPPGEDGRVPE	480
Db	430	AHSASTLDEATPLTWNQSPFLTILQSTNTHHTQSSSSSSDGLFRSPRPAHSLPPGEDGRVPE	489
Qy	481	YVDFAEFYRLWSVDHGQSQSVVTAP 504	
Db	490	YVDFAEFYRLWSVDHGQSQSVVTAP 513	
RESULT 5			
US-10-123-427-6			
; Sequence 6, Application US/10123427			
; Patent No. US20020119525A1			
; GENERAL INFORMATION:			
; APPLICANT: MATSUMOTO, Kunihiro			
; NISHIDA, Eisuke			
; TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR			
; NUMBER OF SEQUENCES: 8			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Foley & Lardner			
; STREET: 3000 K Street, N.W., Suite 500			
; CITY: Washington			
; STATE: D.C.			
; COUNTRY: USA			
; ZIP: 20007-5109			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: Patent In Release #1.0, Version #1.30			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/10/123,427			
; FILING DATE: 17-APR-2002			
; CLASSIFICATION: <Unknown>			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: US/09/406,854			
; FILING DATE: <Unknown>			
; APPLICATION NUMBER: US/08/752,891			
; FILING DATE: 20-NOV-1996			
; APPLICATION NUMBER: JP 8-300856			
; FILING DATE: 28-OCT-1996			
; APPLICATION NUMBER: JP 8-126282			
; FILING DATE: 24-APR-1996			
; ATTORNEY/AGENT INFORMATION:			
; NAME: BENT, Stephen A.			
; REGISTRATION NUMBER: 29,768			
; REFERENCE/DOCKET NUMBER: 17981/111			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: (202) 672-5300			
; TELEFAX: (202) 672-5399			

```
TELEX: 904136
;
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 504 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-123-427-6

Query Match      99.8%; Score 2575; DB 12; Length 504;
Best Local Similarity 99.8%; Pred. No. 1.2e-193;
Matches 503; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAAQRSLIQSEQPSWTDLPLCHLSGVGSASNSYSADGKGSHHPEDSWLKFRSEN 60
    |||
DB 1 MAAQRSLIQSEQPSWTDLPLCHLSGVGSASNSYSADGKGSHHPEDSWLKFRSEN 60
    |||

QY 61 NCFLYGVFNQGNRYTNVAORLSAELLGQLNMAHADVRVYLQAFDVVERSPLES 120
    |||
DB 61 NCFLYGVFNQGNRYTNVAORLSAELLGQLNMAHADVRVYLQAFDVVERSPLES 120
    |||

QY 121 IDDLALAEKASLSQSLPEGVPOHLPPOYOKILERLKTTEREISGMAVAVLNNKLYV 180
    |||
DB 121 IDDLALAEKASLSQSLPEGVPOHLPPOYOKILERLKTTEREISGMAVAVLNNKLYV 180
    |||

QY 181 ANVTNRALLCKSTVDGLQVTLNVDHTTENEDLFRLSQLDAGKIKQVGIIIGQEST 240
    |||
DB 181 ANVTNRALLCKSTVDGLQVTLNVDHTTENEDLFRLSQLDAGKIKQVGIIIGQEST 240
    |||

QY 241 RRTGDKVYKGYTIDLLSAASKPIIARPEIHGAQPLDGVGVFLVMSGGYKRLAEAH 300
    |||
DB 241 RRTGDKVYKGYTIDLLSAASKPIIARPEIHGAQPLDGVGVFLVMSGGYKRLAEAH 300
    |||

QY 301 GPGQANQEIAMIDTEFAKQSLDAVAQAIVDRVKRIHSDTFASGGERARFCPRHEDMTL 360
    |||
DB 301 GPGQANQEIAMIDTEFAKQSLDAVAQAIVDRVKRIHSDTFASGGERARFCPRHEDMTL 360
    |||

QY 361 LVNRFQYPLIGENSQPTPSAPAAGKRVYVSYPYSSAQSTSTKTSVTLSLWPSQGMVNG 420
    |||
DB 361 LVNRFQYPLIGENSQPTPSAPAAGKRVYVSYPYSSAQSTSTKTSVTLSLWPSQGMVNG 420
    |||

QY 421 AHSASTLDEATPTLTNOSTLTLTQSTNHTQSSSSSDGGLFRSRPAHSLPGEGRVREP 480
    |||
DB 421 AHSASTLDEATPTLTNOSTLTLTQSTNHTQSSSSSDGGLFRSRPAHSLPGEGRVREP 480
    |||

QY 481 YVDFAEFYRLMSVDHGEGSVVTAP 504
    |||
DB 481 YVDFAEFYRLMSVDHGEGSVVTAP 504
    |||

RESULT 6
US-09-925-300-1270
; Sequence 1270, Application US/09925300
; Patent No. US20020351681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: Patencin Ver. 2.0
; SEQ ID NO 1270
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
```

```
NAME/KEY: SITE
; LOCATION: (38)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-925-300-1270

Query Match      15.9%; Score 409; DB 10; Length 84;
Best Local Similarity 97.5%; Pred. No. 1.5e-25;
Matches 78; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 425 STUDEATPTLTNOSPPLTLTGSTNHTHTQSSSSSDGGLFRSRPAHSLPGEGRVREPVD 484
    :|||
DB 5 ATLDEATPTLTNOSPPLTLTGSTNHTHTQSSSSSDGGLFRSRPAHSLPGEGRVREPVD 64
    :|||

QY 485 AEFYRLMSVDHGEGSVVTAP 504
    |||
DB 65 AEFYRLMSVDHGEGSVVTAP 84
    |||

RESULT 7
US-09-864-761-34065
; Sequence 34065, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 4917
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34065
; LENGTH: 70
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO Z83845.14
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.77
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.73
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.78
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.78
; OTHER INFORMATION: EST_HUMAN HIT: BE746542.1, EVALUE 5.00e-35
; OTHER INFORMATION: SWISSPROT HIT: Q15750, EVALUE 4.00e-36
US-09-864-761-34065

Query Match      14.1%; Score 365; DB 10; Length 70;
Best Local Similarity 100.0%; Pred. No. 3.2e-22;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 312 MIDTEFAKQTSLDVAQAQVDRVKRIHSDTFASGGERARFCPRHEDMTLLVRNFGYPLGE 371
Db 1 MIDTEFAKQTSLDVAQAQVDRVKRIHSDTFASGGERARFCPRHEDMTLLVRNFGYPLGE 60

Qy 372 MSQPTSPAP 381
Db 61 MSQPTSPAP 70

RESULT 8
US-09-864-761-34067
; Sequence 34067, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aesomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
```

```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34067
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO Z83845.14
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 10
; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 5.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.7
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 8.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 8.9
; OTHER INFORMATION: SWISSPROT HIT: Q15750, EVALUE 7.00e-24
; OTHER INFORMATION: EST_HUMAN HIT: AL118967.1, EVALUE 9.00e-23
US-09-864-761-34067

Query Match      10.1%; Score 261; DB 10; Length 51;
Best Local Similarity 100.0%; Pred. No. 2.7e-14;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 58 SENNCFLYGVFNGYDGNRVTFNFAQRLSABLLLGQNAEHAEDVRRVLLQ 108
Db 1 SENNCFLYGVFNGYDGNRVTFNFAQRLSABLLLGQNAEHAEDVRRVLLQ 51

RESULT 9
US-09-860-351-4
; Sequence 4, Application US/09860351
; Patent No. US20020077463A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 16105, A NOVEL PROTEIN HUMAN PHOSPHATASE
; TITLE OF INVENTION: AND USES THEREFOR
; FILE REFERENCE: 38155-20013.00
; CURRENT APPLICATION NUMBER: US/09/860,351
; CURRENT FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/205,260
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus amino acid
US-09-860-351-4

Query Match      7.9%; Score 205; DB 10; Length 274;
Best Local Similarity 24.3%; Pred. No. 7.2e-09;
Matches 74; Conservative 53; Mismatches 103; Indels 74; Gaps 12;

Qy 34 NRSYSADGKGTEHPEDSWLKFRRSENNCFLYGVFNGYDGNRVTFNFAQRLSABLLLGQ 93
Db 13 NLNSSSSGK-----DSW-----SFFAVDFHGGSQAAYAKGHLHKTILAE- 54

Qy 94 NAEHAEDVRRVLLQAF-DVVERSEFLESIDDALEAKSLQSLPEGVPHQLPPQYQKIL 152
Db 55 -KSFPEGDPWEMKLSDELDALKESFLEADTDELRSAEASAA-----NKVL 99
```



```
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: US 60/304,775
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-973-963-4

Query Match          6.6%; Score 170.5; DB 10; Length 372;
Best Local Similarity 24.7%; Pred. No. 5.5e-06;
Matches 91; Conservative 50; Mismatches 105; Indels 123; Gaps 19;

Qy 21 LPLCHLGVGSASNRYSADGKGTESHPPEDSWLKFRSENNCFLYGVFNGYDGNRVTFV 80
Db 87 IPKISLENVGCSAQ-----IGKRKEN---EDRFDAQLTDEVLYFAVYDGHGPPAAADFC 138
Qy 81 ---AQRLSAELLGQLNAEHAADVRRVLLQAFDVVERSFLESIDDAEAKASLOSQPE 137
Db 139 HTHMEKICMDLLPKEKNLE-----TLLTLAFLE-IDKAFSSHARLSAD--- 180
Qy 138 GVPQHQLPPQYQKILERLKTILEREISGGAMAVAVLLNN-KLYVANVTGNRALLCKSTVD 196
Db 181 -----ATL---LTSGTTATVALLRDGIELVVASVGDSTRAILCRKG-- 217
Qy 197 GLQVTLQNVNDHTTENEDELFRLSQLGLDAGKIKQVGIIIC---GQ-----ESTRIGD 245
Db 218 --KPMKLTIDHTPERKDEKERIKKCG-----GFVAMNSLQOPHVNGLAMTRSIG- 265
Qy 246 YKVKYGYTIDLLSAKSKPIIAEPE-----IHGAQPLDGVGTGFLVLMSEGLYKALEA-- 298
Db 266 -----DLDL-----KTSGVIAEPETKRIKLHHAD-----DSFLVLTDTGDNFMVNSQE 308
Qy 299 -----AHGPGQANQEIAMIDTEFAKQTSLDVAQAQV-----DRVKRIHSD 340
Db 309 ICDFVNCQCHDPNEAAHAV-----TEQAIOYGTEDNSTAVVVPFGAWGKYKNSINFSPSR 363
Qy 341 TFASGGERA 349
Db 364 SFASSGRWA 372

RESULT 13
US-09-973-064-4
; Sequence 4, Application US/09973064
; Patent No. US2002010673A1
; GENERAL INFORMATION:
; APPLICANT: Roch, Jean-Marc
; APPLICANT: Bartel, Paul L.
; APPLICANT: Heichman, Karen
; TITLE OF INVENTION: Protein-Protein Interactions in Neurodegenerative
; FILE REFERENCE: Protein Interactions in ND
; CURRENT APPLICATION NUMBER: US/09/973,064
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/240,790
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-973-064-4

Query Match          6.6%; Score 170.5; DB 10; Length 372;
Best Local Similarity 24.7%; Pred. No. 5.5e-06;
Matches 91; Conservative 50; Mismatches 105; Indels 123; Gaps 19;

Qy 21 LPLCHLGVGSASNRYSADGKGTESHPPEDSWLKFRSENNCFLYGVFNGYDGNRVTFV 80
Db 87 IPKISLENVGCSAQ-----IGKRKEN---EDRFDAQLTDEVLYFAVYDGHGPPAAADFC 138
Qy 81 ---AQRLSAELLGQLNAEHAADVRRVLLQAFDVVERSFLESIDDAEAKASLOSQPE 137
Db 139 HTHMEKICMDLLPKEKNLE-----TLLTLAFLE-IDKAFSSHARLSAD--- 180
Qy 138 GVPQHQLPPQYQKILERLKTILEREISGGAMAVAVLLNN-KLYVANVTGNRALLCKSTVD 196
Db 181 -----ATL---LTSGTTATVALLRDGIELVVASVGDSTRAILCRKG-- 217
Qy 197 GLQVTLQNVNDHTTENEDELFRLSQLGLDAGKIKQVGIIIC---GQ-----ESTRIGD 245
Db 218 --KPMKLTIDHTPERKDEKERIKKCG-----GFVAMNSLQOPHVNGLAMTRSIG- 265
Qy 246 YKVKYGYTIDLLSAKSKPIIAEPE-----IHGAQPLDGVGTGFLVLMSEGLYKALEA-- 298
Db 266 -----DLDL-----KTSGVIAEPETKRIKLHHAD-----DSFLVLTDTGDNFMVNSQE 308
Qy 299 -----AHGPGQANQEIAMIDTEFAKQTSLDVAQAQV-----DRVKRIHSD 340
Db 309 ICDFVNCQCHDPNEAAHAV-----TEQAIOYGTEDNSTAVVVPFGAWGKYKNSINFSPSR 363
Qy 341 TFASGGERA 349
Db 364 SFASSGRWA 372

RESULT 14
US-09-973-077-4
; Sequence 4, Application US/09973077
; Patent No. US20020114799A1
; GENERAL INFORMATION:
; APPLICANT: Roch, Jean-Marc
; APPLICANT: Bartel, Paul L.
; APPLICANT: Heichman, Karen
; TITLE OF INVENTION: Protein-Protein Interactions in Neurodegenerative
; FILE REFERENCE: Protein Interactions in ND
; CURRENT APPLICATION NUMBER: US/09/973,077
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/240,790
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-973-077-4

Query Match          6.6%; Score 170.5; DB 10; Length 372;
Best Local Similarity 24.7%; Pred. No. 5.5e-06;
Matches 91; Conservative 50; Mismatches 105; Indels 123; Gaps 19;

Qy 21 LPLCHLGVGSASNRYSADGKGTESHPPEDSWLKFRSENNCFLYGVFNGYDGNRVTFV 80
Db 87 IPKISLENVGCSAQ-----IGKRKEN---EDRFDAQLTDEVLYFAVYDGHGPPAAADFC 138
Qy 81 ---AQRLSAELLGQLNAEHAADVRRVLLQAFDVVERSFLESIDDAEAKASLOSQPE 137
Db 139 HTHMEKICMDLLPKEKNLE-----TLLTLAFLE-IDKAFSSHARLSAD--- 180
Qy 138 GVPQHQLPPQYQKILERLKTILEREISGGAMAVAVLLNN-KLYVANVTGNRALLCKSTVD 196
Db 181 -----ATL---LTSGTTATVALLRDGIELVVASVGDSTRAILCRKG-- 217
Qy 197 GLQVTLQNVNDHTTENEDELFRLSQLGLDAGKIKQVGIIIC---GQ-----ESTRIGD 245
Db 218 --KPMKLTIDHTPERKDEKERIKKCG-----GFVAMNSLQOPHVNGLAMTRSIG- 265
Qy 246 YKVKYGYTIDLLSAKSKPIIAEPE-----IHGAQPLDGVGTGFLVLMSEGLYKALEA-- 298
Db 266 -----DLDL-----KTSGVIAEPETKRIKLHHAD-----DSFLVLTDTGDNFMVNSQE 308
```

```

QY 299 -----AHGGGQANQELAMIDPEFAKQTSLDVAQAV-----DRKRIHSD 340
Db 309 ICDEVNQCNDENEAHAV-----TEQALQYGTEDNSTAVVVPFGAMGKYKNSEINFSPR 363
QY 341 TFAAGGERA 349
Db 364 SFASGRWA 372

```

RESULT 15

```

US-09-973-063-4
: Sequence 4, Application US/09973063
: Patent No. US20020151199A1
: GENERAL INFORMATION:
: APPLICANT: Koch, Jean-Marc
: APPLICANT: Bartel, Paul L.
: APPLICANT: Hachman, Karen
: TITLE OF INVENTION: Protein-Protein Interactions in Neurodegenerative
: TITLE OF INVENTION: Diseases
: FILE REFERENCE: Protein Interactions in ND
: CURRENT APPLICATION NUMBER: US/09/973,063
: CURRENT FILING DATE: 2001-10-10
: PRIOR APPLICATION NUMBER: US 60/240,790
: PRIOR FILING DATE: 2000-10-17
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 4
: LENGTH: 372
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-973-063-4

```

Query Match	6.6%	Score 170.5	DB 10	Length 372
Best Local Similarity	24.7%	Pred. No. 5.5e-06		
Matches 91	Conservative 50	Mismatches 105	Indels 123	Gaps 19

Search completed: December 9, 2002, 23:04:26
Job time : 101.986 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 9, 2002, 22:58:34 ; Search time 3162.68 Seconds
(without alignments)
4637.778 Million cell updates/sec

Title: US-09-830-144-4
Perfect score: 2580
Sequence: 1 MAQRSLQEQOPSMTDD.....AEFYRLMSVDHGQSVVTAP 504

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlp
-Q/cgn2_1/USPTO.spool/US09830144/runat_04122002_141353_2252/app.query.fasta_1.1422
-DB=GenEmbl -QFMT=fastcap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -NORM=ext -HEAPSIZ=500 -MINLEN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09830144 @CNG 1 1 5173 @runat 04122002 141353 2252 -NCPU=6 -ICPU=3
-NO_XUPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sv.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*

- 29: em.vi.*
- 30: em.htg_hum.*
- 31: em.htg_inv.*
- 32: em.htg_other.*
- 33: em.htg_mus.*
- 34: em.htg_pln.*
- 35: em.htg_rnd.*
- 36: em.htg_wam.*
- 37: em.htg_vrt.*
- 38: em.sv.*
- 39: em.htgo_hum.*
- 40: em.htgo_mus.*
- 41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2580	100.0	1515	6	E31041 Method for
2	2580	100.0	1560	6	AR058299 Sequence
3	2580	100.0	1560	6	AR088273 Sequence
4	2580	100.0	1560	6	AR116881 Sequence
5	2580	100.0	1560	6	E14752 Human mRNA
6	2580	100.0	3096	9	U49928 Homo sapien
7	2575	99.8	1560	6	AR058302 Sequence
8	2575	99.8	1560	6	AR088276 Sequence
9	2575	99.8	1560	6	AR116884 Sequence
10	2512	97.4	2944	10	BC027054
11	1924	74.6	1926	5	U92031 Xenopus lae
12	547.5	21.2	69660	9	Z83845 Human DNA s
C	457	17.7	176685	2	AC127784 Rattus no
	414	16.0	176685	2	AC127784 Rattus no
14	414	16.0	141498	2	AC127924 Rattus no
15	387	15.0	141498	2	AC127924 Rattus no
C	369	14.3	141498	2	AC127924 Rattus no
	258.5	10.0	1161	3	AF145375 Caenorhab
17	208	8.1	1479	8	AB083482 Mesembrya
18	205.5	8.0	967	8	AY133761 Arabidops
19	205.5	8.0	1453	8	AY080658 Arabidops
20	203.5	7.9	1824	6	AX319337 Sequence
21	202	7.8	1851	8	AY086406 Arabidops
22	198	7.7	1463	8	AB079671 Arabidops
23	198	7.7	1522	8	AY050873 Arabidops
24	198	7.6	1099	8	AY091209 Arabidops
25	196.5	7.5	849	8	AF370608 Arabidops
26	194.5	7.5	31332	8	AP005244 Oryza sat
C	194	7.5	852	8	AY133563 Arabidops
	193.5	7.5	96221	8	AC051631 Arabidops
29	193.5	7.5	109367	8	AC023673 Genomic s
30	193.5	7.5	1194	8	AY048244 Arabidops
31	192.5	7.4	1345	8	AY080735 Arabidops
32	192	7.4	1422	8	AY085196 Arabidops
33	190.5	7.4	1531	8	AY057611 Arabidops
34	190.5	7.4	1531	8	AY062454 Arabidops
35	190.5	7.4	1902	8	AY062454 Arabidops
36	189	7.3	877	8	AY133737 Arabidops
37	188.5	7.3	1074	8	AY113024 Arabidops
38	188.5	7.3	1479	8	AY087522 Arabidops
39	187.5	7.3	1259	8	AY084874 Arabidops
40	187	7.2	2204	8	AY065026 Arabidops
41	186.5	7.2	1689	8	AY062587 Arabidops
42	186	7.2	1248	8	AF092432 Lotus jap
43	186	7.2	1767	8	AF499718 Theillungi
44	185	7.2	1966	8	AY045819 Arabidops
45	184.5	7.2	1263	8	AY133656 Arabidops

ALIGNMENTS

E31041
LOCUS E31041 1515 bp DNA linear PAT 18-JUN-2001
DEFINITION Method for screening substance inhibiting binding to XIAP.
ACCESSION E31041
VERSION E31041.1 GI:13017306
KEYWORDS JP 199326328-A/1.
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1515)
AUTHORS Kunhiro, M.
TITLE Method for screening substance inhibiting binding to XIAP
JOURNAL Parent: JP 199326328-A 1 26-NOV-1999;
KUNHIRO MATSUMOTO
COMMENT OS Unidentified
PN JP 199326328-A/1
PD 26-NOV-1999
PF 13-MAY-1998 JP 1998130378
PR
PI KUNHIRO MATSUMOTO
PC GOIN33/536, GOIN33/536//C12N15/09, C12P21/08, A61K37/02,
PC A61K37/02, PC A61K37/02, A61K37/24, C12N15/00
PC A61K45/00, A61K45/00, A61K45/00, C07K7/06, C07K7/08, C07K14/47, PC
GOIN33/536,
PC GOIN33/536, GOIN33/536//C12N15/09, C12P21/08, A61K37/02,
A61K37/02, PC A61K37/02, A61K37/24, C12N15/00
CC Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers
FT CDS 1..1512.
FEATURES
source 1..1515 Location/Qualifiers
BASE COUNT 323 a 457 c 463 g 272 t
ORIGIN
Alignment Scores:
Pred. No.: 3,856-151 Length: 1515
Score: 2580.00 Matches: 504
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-09-830-144-4 (1-504) x E31041 (1-1515)
QY 1 MetAlaAlaGlnArgSerLeuLeuGlnInserGlnGlnInProSerTyrThrAspAsp 20
DB 1 ATGGCGGCGGAG 60
QY 21 LeuProLeuCyHisLeuSerGlyValGlySerAlaSerAsnArgSerTyrSerAlaAsp 40
DB 61 CTGGCTCTCTGGCACCCTCTGTGGGGTGTGGCTCAGCCCTCAACCGCAGCTTACTGTGGAT 120
QY 41 GlyLysGlyThrGlnSerHisProProGlnAspSerTyrPheLysSerGlnAsn 60
DB 121 GCGAAGGCGACATGAGAGCCAGCCGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
QY 61 AsnCySPheLeuTyrGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnPheVal 80
DB 181 AACTGCTTCTCTGTATGGGGCTCTTCAACGGCTATGATGCAACCGAGTACCAACTTGTG 240
QY 81 AlaGlnArgLeuSerAlaGlnLeuLeuGlnGlnLeuAsnAlaGlnHisAlaGlnAla 100
DB 241 GCCCGAGCGGTGTCCGCGAGAGACTCTGTGGGCGCAGTGAATGCCAGCAGCAGAGCC 300
QY 101 AspValArgValLeuLeuGlnAlaPheAspValValGlnArgSerPheLeuGlnSer 120
DB 301 GATGTGCGCGGTGTGTGCTGCTCAGGCCCTTCATGTGTGTGGAGAGAGAGAGAGAGAG 360
QY 121 IleAspAspAlaLeuAlaGlnLysAlaSerLeuGlnInserGlnLeuProGlnGlyValPro 140

DB 361 ATTACACAGCGCTTGGCTGGAAGGCGAAGCCTCAGTCCGAAATTCGACAGGAGAGTCCCT 420
QY 141 GlnHisGlnLeuProProGlnTyrGlnLysIleLeuGlnArgLeuLysThrLeuGlnArg 160
DB 421 CAGCACAGCTGCTCTCTCTGATCATCGAATGATCTTGAGAGACTTGAAGAGCTTAAGAGAG 480
QY 161 GluIleSerGlyGlyAlaMetAlaValAlaValLeuLeuAsnAsnLysLeuTyrVal 180
DB 481 GAAATTCGAGAGGAGGCGCATGCGGTGTGGCGGTCTTCTCAACAACAGCTTACGTC 540
QY 181 AlaAsnValGlyThrAsnAlaLeuLeuCyLysSerThrValAspGlyLeuGlnVal 200
DB 541 GCCAATGTCGTACCAACCGTGCATTTATGCAAAATGCAAGTGGATGGTGCAGGTG 600
QY 201 ThrGlnLeuAsnValAspHisThrThrGlnAsnGlnAspGlnLeuPheArgLeuSerGln 220
DB 601 ACACAGCTGAACGTGAGACCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 221 LeuGlyLeuAspAlaGlyLysIleLysGlnValGlyIleIleCySerGlyGlnInserThr 240
DB 661 CTGGGCTTGGATGCTGGAAAGATCAAGCAGGTGGGATCATCTGTGGCAGAGAGAGACAC 720
QY 241 ArgArgIleGlyAspTyrLysValLysTyrGlyTyrThrAspIleAspLeuSerAla 260
DB 721 CGGGGATCGGGGATTAACAAGTTAATGCTTACACGAGCAATTCCTTCAGCGCT 780
QY 261 AlaLysSerLysProIleIleAlaGlnProGlnIleHisGlyValAlaGlnProLeuAspGly 280
DB 781 GCCAAGTCCAAACCATCATCGAGAGCCAGAAATCATGGGCGACAGCGCTGAGAGAG 840
QY 281 ValThrGlyPheLeuValLeuMetSerGlnGlyLeuTyrLysAlaLeuGlnAlaHis 300
DB 841 GTGACGGGCTTCTGTGGTGTGATGCGAGAGGTTGTACAGGCCCTTACAGGAGAGCCAT 900
QY 301 GlyProGlyGlnAlaAsnGlnLysIleAlaMetIleAspThrGlnPheAlaLysGln 320
DB 901 GGGCTGTGGGAGGCGCAACGAGAGATTGCTGCATGATTAAGACATGATTCGCAAGCAG 960
QY 321 ThrSerLeuAspAlaValAlaGlnAlaValAlaAspArgValLysArgIleHisSerAsp 340
DB 961 ACTTCCTGGACCCATATGGCCAGAGCGTCGTGAGACCGGGTGAAGCGCATTCACAGCGAC 1020
QY 341 ThrPheAlaSerGlyGlyGlnArgAlaArgPheCySProArgHisGlnAspMetThrLeu 360
DB 1021 ACCCTGCCAGTGTGGGAGCGTGCAGGTTCTGCCCCGCGCAGAGAGACATGACCTGTG 1080
QY 361 LeuValArgAsnPheGlyTyrProLeuGlnGlyLysSerGlnProThrProSerProAla 380
DB 1081 CTATGAGGAACCTTGTGCTAACCGGCTGGGCGAAATGAGCCAGACCCAGCGCCAGCC 1140
QY 381 ProAlaAlaGlyGlyArgValTyrProValSerValProTyrSerSerAlaGlnSerThr 400
DB 1141 CCACTGCGAGAGAGAGAGATGATACCTGTGTGTGCCATATCTCAGCGCCAGAGACAC 1200
QY 401 SerLysThrSerValThrLeuSerLeuValMetProSerGlnGlyGlnMetValAsnGly 420
DB 1201 AGCMAAGCAGCGTACCTCTCTCTGTCATGCCCTCCAGAGGCGCAGATGGTCAACGGG 1260
QY 421 AlaHisSerAlaSerThrThrLeuAspGlnAlaThrProThrLeuThrAsnGlnSerProThr 440
DB 1261 GCTCAAGAGTCTTCAACCTTGAGCAAGGCAACCCACCTTACCAACCAAGCCCGGACC 1320
QY 441 LeuThrLeuGlnSerThrAsnThrHisThrGlnInserSerSerSerSerAspGlyGly 460
DB 1321 TTAACTTCAGTCCACCAACACGACAGCAGACAGTCCAGCTTCAGCGAGAGCC 1380
QY 461 LeuPheArgSerArgProAlaHisSerLeuProProGlyGlyLysGlyArgValGlnPro 480
DB 1381 CTCTTCCGCTCCCGGCGCGCCACATCGCTCCGCGCTGGCGAGAGAGAGAGAGAGAG 1440
QY 481 TyrValAspPheAlaGlnPheTyrArgLeuTyrPheSerValAspHisGlyGlnInserVal 500

Db 1441 TATGTGGACTTTGCTGAGTTTACCAGCTCTCGAGCGTGGACCATGGAGCAGAGCGTG 1500

QY 501 ValThrAlaPro 504

Db 1501 GTGACAGCACCG 1512

RESULT 2

AR058299

LOCUS AR058299 1560 bp DNA linear PAT 29-SEP-1999

DEFINITION Sequence 1 from patent US 5837819.

ACCESSION AR058299

VERSION AR058299.1 GI:5983876

KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 1560)

AUTHORS Matsumoto,K. and Nishida,E.

TITLE TAB1 protein

JOURNAL Patent: US 5837819-A 1 17-NOV-1999;

FEATURES

source Location/Qualifiers

1..1560

/organism="unknown"

BASE COUNT 332 a 469 c 480 g 279 t

ORIGIN

Alignment Scores:

Pred. No.: 3,99e-151 Length: 1560

Score: 2580.00 Matches: 504

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 6 Gaps: 0

US-09-830-144-4 (1-504) x AR058299 (1-1560)

QY 1 MetAlaAlaGlnArgArgSerLeuLeuGlnSerGluGlnGlnProSerTrpThrAspAsp 20

Db 30 ATGGCGCGCAGAGGAGGAGCTTGCTGCAGAGTGAGCAGCAGCCAAAGCTGGACAGATGAC 89

QY 21 LeuProLeuCysHisLeuSerGlyValGlySerAlaSerAsnArgSerTyrSerAlaAsp 40

Db 90 CTGCCTCTCTGCCACCTCTCTGGGTTGGCTCAGCCTCCAAACCGCAGCTACTCTGTCTGAT 149

QY 41 GlyValGlyThrGluSerHisProProGluAspSerTrpLeuLysPheArgSerGluAsn 60

Db 150 GGCAAGGGCACTGAGAGCCACCCGCCAGAGCAGCTGGCTCAAGTTCAGGAGTGAGAAC 209

QY 61 AsnCysPheLeuTyrGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnPheVal 80

Db 210 AACTGCTTCTGTATGGGGTCTTCAAGGGCTATGATGGCAACCGAGTGACCAACTTCGTG 269

QY 81 AlaGlnArgLeuSerAlaGluLeuLeuGlyGlnLeuAsnAlaGluHisAlaGluAla 100

Db 270 GCCCAGCGGCTGTCCGACAGCTCTCTGGGCAGCTGAATGCCGACGACGCCGAGGCC 329

QY 101 AspValArgArgValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGluSer 120

Db 330 GATGTGGCGGCTGTGCTGTCAGGCCCTTCGATGTGTGGAGGAGCTTCCTGGAGTCC 389

QY 121 IleAspAspAlaLeuAlaGluLysAlaSerLeuGlnSerGlnLeuProGluGlyValPro 140

Db 390 ATTGACAGCGCTTGGCTGAGAGGCCAAGCCTCCAGTCGCAATTGCCAGAGGGAGTCCCT 449

QY 141 GlnHisGlnLeuProProGlnTyrGlnLysIleLeuGluArgLeuLysThrLeuGluArg 160

Db 450 CAGCACCAAGCTGCCTCTCAGTATCAGAAGATCCTTGAGAGACTCAAGACGTTAGAGAG 509

QY 161 GluIleSerGlyAlaMetAlaValAlaValAlaValLeuLeuAsnLeuLysLeuTyrVal 180

Db 510 GAAATTTGGGAGGGGCATGGCGCTGTGTGGCGGTCTCTTCAACAAACAAGCTCTACGTC 569

QY 181 AlaAsnValGlyThrAsnArgAlaLeuLeuCysLysSerThrValAspGlyLeuGlnVal 200

Db 570 GCCAATGTCGGTACAAACCGTGCACCTTTTATGCAAAATCGACAGTGGATGGTTGCAAGTG 629

QY 201 ThrGlnLeuAsnValAspHisThrThrGluAsnGluAspGluLeuPheArgLeuSerGln 220

Db 630 ACACAGCTGAACGTGGACCAACACACAGAAACAGAGATGAGCTCTTCCGTCTTTGCAG 689

QY 221 LeuGlyLeuAspAlaGlyLysIleLysGlnValGlyIleLysCysGlyGlnGluSerThr 240

Db 690 CTGGCTTGGATGCTGGAAGATCAAGCAGCTGGGATCATCTGTGGCAGGAGAGACACC 749

QY 241 ArgArgIleGlyAspTyrLysValLysTyrGlyTyrThrAspIleAspLeuLeuSerAla 260

Db 750 CGCGGATCGGGATTACAAGGTTAAATATGGCTACACGACATTGACCTTCTCAGCGCT 809

QY 261 AlaLysSerLysProIleIleAlaGluProGluIleHisGlyAlaGlnProLeuAspGly 280

Db 810 GCCAAGTCCAAACCAATCATCGCAGAGCCAGAAATCCATGGGGCAGACCGCTGGATGGG 869

QY 281 ValThrGlyPheLeuValLeuMetSerGluGlyLeuTyrLysAlaLeuGluAlaHis 300

Db 870 GTGACGGCTTCTTGGTGCTGATGTCGAGGGGTTGTACAAGGCCCTAGAGGCAGCCAT 929

QY 301 GlyProGlyGlnAlaAsnGlnGluIleAlaAlaMetIleAspThrGluPheAlaLysGln 320

Db 930 GGGCTGGCAGGCCAAACAGGAGATTGCTCGCATGATTGACACTGAGTTTGCACAAGCAG 989

QY 321 ThrSerLeuAspAlaValAlaGlnAlaValAlaAspArgValLysArgIleHisSerAsp 340

Db 990 ACCTCCCTGGACGAGTGGCCAGGCCGCTGCTGGACCGGGTGAAGCGCATCCACAGCGAC 1049

QY 341 ThrPheAlaSerGlyGlyGluArgAlaArgPheCysProArgHisGluAspMetThrLeu 360

Db 1050 ACCTTCGCCAGTGTGGGAGCGTCCAGGTTCTGCCCGCCGACAGGACATGACCCCTG 1109

QY 361 LeuValArgAsnPheGlyTyrProLeuGlyGluMetSerGlnProThrProSerProAla 380

Db 1110 CTAGTAGGAACTTTGGCTACCCGCTGGCGAAATGAGCCAGCCACACCGAGCCAGCC 1169

QY 381 ProAlaAlaGlyGlyArgValTyrProValSerValProTyrSerSerAlaGlnSerThr 400

Db 1170 CCAGTGCAGAGGAGCAGATGATCCCTGTGTGTGCATACTCCAGCGCCAGAGCACC 1229

QY 401 SerLysThrSerValThrLeuSerLeuValMetProSerGlnGlyGlnMetValAsnGly 420

Db 1230 AGCAAGACCGGTGACCTCTCCCTTGTCTATGCTCCAGGGCCAGATGGTCAACGGG 1289

QY 421 AlaHisSerAlaSerThrLeuAspGluAlaThrProThrLeuThrAsnGlnSerProThr 440

Db 1290 GCTCACAGTGTCTCCACCTCGAGCAAGCCACCCCTCACCAACCAAGCCCGAC 1349

QY 441 LeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSerSerAspGlyGly 460

Db 1350 TTAAACCTGTCAGTCCACCAACGACGACGAGCAGCAGCTCCAGCTCTGACGAGGC 1409

QY 461 LeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGlyArgValGluPro 480

Db 1410 CTCCTCCGCTCCCGGCCCGCCACTCGCTCCCGCTGCGAGGAGCGTCTGTTGAGGCC 1469

QY 481 TyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGlyGluGlnSerVal 500

Db 1470 TATGTGGACTTGTCTGAGTTTATCCGCCCTCTGGAGCGTGGACCATGGCGCAGAGCGTG 1529

QY 501 ValThrAlaPro 504

Db 1530 GTGACAGCACCG 1541

RESULT 3

AR088273

LOCUS AR088273 1560 bp DNA linear PAT 07-SEP-2000

DEFINITION Sequence 1 from patent US 5989862.

ACCESSION AR088273

VERSION AR088273.1 GI:10015036

KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 1560)
TITLE Tabi protein and DNA coding therefor
JOURNAL Patent: US 5989862-A 1 23-NOV-1999;
FEATURES Location/Qualifiers
source 1..1560
/organism="unknown"
BASE COUNT 332 a 469 c 480 g 279 t
ORIGIN
Alignment Scores:
Pred. No.: 3,99e-151 Length: 1560
Score: 2580.00 Matches: 504
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-09-830-144-4 (1-504) X AR088273 (1-1560)
QY 1 MetAlAlAGlAArgArSerLeuLeuGlnGlnProSerThrAspAsp 20
DB 30 ATGGCGGCGCAGAGGAGCTTGCGAGAGACAGACCCAGCTGACAGATGAC 89
QY 21 LeuProLeuCyHisLeuSerGlyValGlySerAlaSerAsnArgSerTyrSerAlaAsp 40
DB 90 CTCCCTCTGCGACCTCTCTGGGGTGGCTCAGCCTCCACCGACGACTCTCTGAT 149
QY 41 GlyVSGlyThrGluSerHisProProGluAspSerTyrPheArgSerGluAsn 60
DB 150 GGGAGGGGACATGAGAGCCAGCCGAGAGAGACAGCTGCTCAAGTTCAAGAGTGAAC 209
QY 61 AsnCySPheLeuTyrGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnPheVal 80
DB 210 AACTGCTTCTGTATGGGCTTTCACAGGCTATGATGGCACCGAGTACCAACTTCGTG 269
QY 81 AlaGlnArgLeuSerAlaGluLeuLeuLeuGlyGlnLeuAsnAlaGlnHisAlaGlnAla 100
DB 270 GCCGACGGGCTGTCCGACAGCTCTCGTGGGCCAGCTGAAGCCGAGACCCGAGGCC 329
QY 101 AspValArgArgValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGlnSer 120
DB 330 GATGTGCGGCGTGTGCTGTGCGAGGCTTCGATGTGTGAGAGAGAGCTTCTCGAGTCC 389
QY 121 IleAspAspAlaLeuAlaGluValAspLeuGlnSerGlnLeuProGluGlyValPro 140
DB 390 ATTGAGAGAGCTTGGCTGAGAGGCAAGCTTCAGTCCCAATTCGACAGGGAGTCCCT 449
QY 141 GlnHisGlnLeuProProGlnTyrGlnTyrSileuGlnArgLeuLeuThrLeuGlnArg 160
DB 450 CAGCACACAGCTGCTCTCAGTATCAGAAGATCCTTGAGAGCTCAAGACCTGAGAGG 509
QY 161 GluIleSerGlyValAlaMetAlaValAlaValLeuLeuAsnAsnLysLeuTyrVal 180
DB 510 GAAATTTTCGGAGGGGCCATGTGGCTTGTGGCTCTTCACAAACAAGCTTACGTC 569
QY 181 AlaAsnValAlaGlyThrAsnArgAlaLeuLeuCySylSerSerThrValAspGlyLeuGlnAla 200
DB 570 GCCAATGTCGTAACAAACCGTCCACTTTTATGCATAATGCAGTGGATGGGTGAGGTG 629
QY 201 ThrGlnLeuAsnValAspHisThrThrGlnGlnGlnAspGluLeuPheArgLeuSerGln 220
DB 630 ACACAGCTGAACGTGCACACACACACAGAGAAAGAGATGAGCTTTCGCTTTCGAG 689
QY 221 LeuGlyLeuAspAlaGlyValSileuSglnValGlyIleIleCySglnGlnGlnSerThr 240
DB 690 CTGGGCTTGGATGCTGGAAGATCAAGCAGGTGGGATCATCTGTGGGCGAGAGGACCC 749
QY 241 ArgArgIleGlyAspTyrLysValLysTyrGlyTyrThrAspIleAspLeuLeuSerAla 260

DB 750 CGCGGATCGGGGATTAACAAGTTAATGCTACACGAGCATTCCTTCAGGCT 809
QY 261 AlaLysSerLysProIleIleAlaGluProGluIleHisGlyValGlnProLeuAspGly 280
DB 810 GCCAATGCCAAACCAATCATCGCAGAGCCAGAAATCCATGGGCGACGCCGTGATGGG 869
QY 281 ValThrGlyPheLeuValLeuMetSerGluGlyLeuTyrLysAlaLeuGlnAlaHis 300
DB 870 GTGACGGGCTTCTTGGTGTGATGTGGAGGGTTGTACAAAGCCCTAAGAGGAGCCCT 929
QY 301 GlyProGlyGlnAlaAsnGlnGluIleAlaAlaMetIleAspThrGluPheAlaLysGln 320
DB 930 GGGCGTGGGAGGCCAACCGAGAGATTGCTGCATGATTGACATGATTGGCCAGACAG 989
QY 321 ThrSerLeuAspAlaValAlaGlnAlaValAlaAspArgValLysArgIleHisSerAsp 340
DB 990 ACTTCCTGGAGCGCAGTGGCCAGGCCGTGTGACCGGGTGAAGCGCATCCACCGAC 1049
QY 341 ThrPheAlaSerGlyGlyGluArgAlaArgPheCySPheArgHisGluAspMetThrLeu 360
DB 1050 ACTTCGCCAGTGGTGGGAGCGGCGAGGTTCTGCCCCCGGACAGAGACATGACCTG 1109
QY 361 LeuValArgAsnPheGlyTyrProLeuGlyGluMetSerGlnProThrProSerProAla 380
DB 1110 CTAGTGAAGAACTTTGGCTACCCGCTGGGCGAAATGAGCCAGCCACCGAGCCAGCC 1169
QY 381 ProAlaAlaGlyArgValTyrProValSerValProTyrSerSerAlaGlnSerThr 400
DB 1170 CCAAGTCCAGAGAGAGAGAGTACCTGTGTCTGTGCCATCTCCAGCGCCAGAGACACC 1229
QY 401 SerLysThrSerValThrLeuSerLeuValMetProSerGlnGlyGlnMetValAsnGly 420
DB 1230 AGCAAGACACAGCGTACCTCTCTTGTGCATGCTCCAGAGGCCAGATGGTCAACGG 1289
QY 421 AlaHisSerLysSerThrLeuAspGluAlaThrProThrLeuThrAsnGlnSerProThr 440
DB 1290 GCTCACAGTCTTCCACCTCGAGAGAGCCACCCACCTCACCAACCAAGCCCGAGCC 1349
QY 441 LeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSerAspGlyGly 460
DB 1350 TTAAACCTGCACTCCACCAACACGACACGACAGAGACAGAGCTTCAGCGAGGCC 1409
QY 461 LeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGlyArgValGluPro 480
DB 1410 CTCTTCGCTCCCGGCCCGCCACTCGCTCCGCTGGGAGAGACGCTGTGTGAGGCC 1469
QY 481 TyrValAspPheAlaGluPheTyrArgLeuTyrPheSerValAspHisGlyGlnSerVal 500
DB 1470 TATGTGACTTGTCTGAGTTTAAACGCTCTGTGAGCGTGAACCATGGCGAGAGACGTG 1529
QY 501 ValThrAlaPro 504
DB 1530 GTGACAGCACCG 1541
RESULT 4
AR116881 1560 bp DNA linear PAT 16-MAY-2001
LOCUS AR116881
DEFINITION Sequence 1 from patent US 6140042.
ACCESSION AR116881
VERSION AR116881.1 GI:14097787
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 1560)
TITLE Tabi protein and DNA coding therefor
JOURNAL Patent: US 6140042-A 1 31-OCT-2000;
FEATURES Location/Qualifiers
source 1..1560
/organism="unknown"
BASE COUNT 332 a 469 c 480 g 279 t

ORIGIN

Alignment Scores: 3.99e-151 Length: 1560
 Pred. No.: 2580.00 Matches: 504
 Score: 100.00% Conservative: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 100.00% Gaps: 0
 DB: 6

US-09-830-144-4 (1-504) x AR116881 (1-1560)

Qy 1 MetAlaAlaGlnArgSerLeuLeuGlnSerGluGlnGlnProSerTrpThrAspAsp 20
 Db 30 ATGGCGCGCAGAGGAGGAGCTTGCTGCAGAGTGAGCAGCAGCCAAAGCTGGACATGAC 89
 Qy 21 LeuProLeuCysHisLeuSerGlyValGlySerAlaSerAsnArgSerTyrSerAlaAsp 40
 Db 90 CTGCCTCTCTGCCACCTCTCTGGGGTGGCTCAGCCCTCCAACCGCAGCTTCTCTGTAT 149
 Qy 41 GlyLysGlyThrGluSerHisProProGluAspSerTyrLeuLysPheArgSerGluAsn 60
 Db 150 GGCAAGGCACTGAGAGCACCAGCCGAGAGCAGCTGGCTCAAGTTCAAGGAGTGAGAAC 209
 Qy 61 AsnCysPheLeuTyrGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnPheVal 80
 Db 210 AACTGCTTCTGTATGGGGTCTTCAACGGCTATGATGGCAACCGAGTGACCAACTTCGTG 269
 Qy 81 AlaGlnArgLeuSerAlaGluLeuLeuLeuGlyGlnLeuAsnAlaGluHisAlaGluAla 100
 Db 270 GCCCAGCGGCTGTCGAGAGCTCTCTGGCCAGCTGAATGCCGAGCAGCCGAGGCC 329
 Qy 101 AspValArgArgValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGluSer 120
 Db 330 GATGTGGGGCTGTGCTGTCAGAGGCTTCGATGTGGTGGAGAGAGCTTCCTGGAGTCC 389
 Qy 121 IleAspAspAlaLeuAlaGluLysAlaSerLeuGlnSerGlnLeuProGluGlyValPro 140
 Db 390 ATTGACGACCCCTTGGCTGAGAGGCAAGCCCTCAGTCGCAATTGCCAGAGGAGTCCCT 449
 Qy 141 GlnHisGlnLeuProProGlnTyrGlnLysIleLeuGluArgLeuLysThrLeuGluArg 160
 Db 450 CAGCACCAGCTGCTCTCAGTATCAGAAATCCTTTGAGAGACTCAAGACGTTAGAGAGG 509
 Qy 161 GluIleSerGlyGlyAlaMetAlaValAlaValLeuLeuAsnAsnLysLeuTyrVal 180
 Db 510 GAAATTTGGGAGGGGCCATGGCGCTTGTGGCGCTCTTCTCAACACCAAGCTCTACGTC 569
 Qy 181 AlaAsnValGlyThrAsnArgAlaLeuLeuCysLysSerThrValAspGlyLeuGlnVal 200
 Db 570 GCCAATGTCGGTACAAACCGTGACATTTTATGCAAAATCGACAGTGATGGTTGCAGGTG 629
 Qy 201 ThrGlnLeuAsnValAspHisThrThrGluAsnGluAspGluLeuPheArgLeuSerGln 220
 Db 630 ACACAGCTGAACGTGGACCAACACACAGAGACGAGATGAGCTCTTCGCTCTTTCCGACG 689
 Qy 221 LeuGlyLeuAspAlaGlyLysIleLysGlnValGlyIleLysCysGlyGlnGluSerThr 240
 Db 690 CTGGGCTTGGATGCTGGAAGATCAACAGGTGGGGATCATCTGTGGGAGGAGACACC 749
 Qy 241 ArgArgIleGlyAspTyrLysValLysTyrGlyTyrThrAspIleAspLeuLeuSerAla 260
 Db 750 CGGCGGATCGGGATTACAAGTTAAATATGGCTACACGAGCATTTGACCTTCTCAGCGCT 809
 Qy 261 AlalysSerLysProIleAlaGluProGluIleHisGlyAlaGlnProLeuLeuAspGly 280
 Db 810 GCCAAGTCCAAACCAATCATCGACAGCCAGAAATCCATGGGACACGCGCTGGATGGG 869
 Qy 281 ValThrGlyPheLeuValLeuMetSerGluGlyLeuTyrLysAlaLeuGluAlaHis 300
 Db 870 GTGACGGGCTTCTTGGTGTGCTGATCGGAGGGGTGTACAAAGGCCCTTAGAGCGGCCAT 929
 Qy 301 GlyProGlyGlnAlaAsnGlnGluIleAlaAlaMetIleAspThrGluPheAlaLysGln 320

Db 930 GGGCCTGGCAGGCCAACACAGGAGATTGCTGCGATGATTGACACTGATTGTCACAGCAG 989
 Qy 321 ThrSerLeuAspAlaValAlaGlnAlaValAlaAspArgValLysArgIleHisSerAsp 340
 Db 990 ACCTCCCTGGACGAGTGCCAGGCGCTGCTGGAACGGGTGAAGCGCATCCACAGGCAC 1049
 Qy 341 ThrPheAlaSerGlyGlyGluArgAlaAraPheCysProArgHisGluAspMetThrLeu 360
 Db 1050 ACCTTCGCCAGTGTGGGAGCGTGCAGGTTCTGCCCGCCGACGAGGACATGACCCCTG 1109
 Qy 361 LeuValArgAsnPheGlyTyrProLeuGlyGluMetSerGlnProThrProSerProAla 380
 Db 1110 CTAGTGAGGAACCTTTGGCTACCCGCTGGGCGAAATGAGCCAGCCACACCGAGCCAGCC 1169
 Qy 381 ProAlaAlaGlyArgValTyrProValSerValProTyrSerSerAlaGlnSerThr 400
 Db 1170 CCAGCTCGAGAGGACGAGTGATCCCTGTGTGTGTCATACTCCAGCGCCAGAGCACC 1229
 Qy 401 SerLysThrSerValThrLeuSerLeuValMetProSerGlnGlyGlnMetValAsnGly 420
 Db 1230 AGCAAGACCAGCGTGACCTCTCCCTTGTTCATGCTCCCTCCAGGGCCAGATGGTCAACGGG 1289
 Qy 421 AlaHisSerAlaSerThrLeuAspGluAlaThrProThrLeuThrAsnGlnSerProThr 440
 Db 1290 GCTCACAGTGTCTCACCTGGACGAAGCCACCCCTCCACCAACCAAGCCCGACCC 1349
 Qy 441 LeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSerSerSerGlyGly 460
 Db 1350 TTAACCTTGCAGTCCACCAACACGACGACGAGCAGCAGCTCCAGCTCTGACGGAGCC 1409
 Qy 461 LeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGlyArgValGluPro 480
 Db 1410 CTCITCCGCTCCCGCGCCGCCACTCGCTCCCGCTGGCAGGAGCGCTGCTGTTGAGCCC 1469
 Qy 481 TyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGlyGluGlnSerVal 500
 Db 1470 TATGTGGACTTGTGTGAGTTTACCCTCTGAGCGCTGAGACCTGAGCAGCAGGAGCGTG 1529
 Qy 501 ValThrAlaPro 504
 Db 1530 GTGACAGCACCG 1541

RESULT 5
 E14752 Human mRNA for the TAB1 protein. 1560 bp DNA linear PAT 28-JUL-1999
 LOCUS E14752
 DEFINITION Human mRNA for the TAB1 protein.
 ACCESSION E14752
 VERSION E14752.1 GI:5709435
 KEYWORDS JP 1998004976-A/1.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1560)
 AUTHORS Matsumoto, K. and Nishida, E.
 TITLE TAB1 PROTEIN AND DNA CODING THE SAME
 JOURNAL Patent: JP 1998004976-A 1 13-JAN-1998;
 UENO NAOTO
 COMMENT OS Homo sapiens (human)
 PN JP 1998004976-A/1
 PD 13-JAN-1998
 PF 28-OCT-1996 JP 1996300856
 PR 24-APR-1996 JP 96P 126282
 PI MATSUMOTO KUNIHICO, NISHIDA EISUKE
 PC C12N15/09, C07H21/04, C07K14/47, C07K19/00, C12N1/19, C12N1/21, PC
 C12N5/10,
 PC C12P21/02, G01N33/53, (C12N1/19, C12R1:865), (C12N1/21, C12R1:19),
 PC (C12N5/10,
 PC C12R1:91), (C12P21/02, C12R1:865), (C12P21/02, C12R1:91); CC
 strandedness: Double;
 CC topology: Linear;
 FH Key Location/Qualifiers

FH source 1..1560
 /organism='Homo sapiens'
 /tissue='kidney'
 /product='the TMB1 protein'
 /note='replace a'.
 FEATURES
 source 1..1560
 /organism='Homo sapiens'
 /db_xref='taxon:9606'
 BASE COUNT 332 a 469 c 480 g 279 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 3.99e-151 Length: 1560
 Score: 2580.00 Matches: 504
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0
 US-09-830-144-4 (1-504) X E14752 (1-1560)
 QY 1 MetAlAlAGLInrGArSerLeuLeuGInGInProSerTTPThrAspAsp 20
 Db 30 ATGGCGGCGCAGAGGAGGAGCTTGCTGACAGTGAACACACGACGACATGAC 89
 QY 21 LeuProLeuCySerHisLeuSerGlyValGlySerAlaSerAsnArgSerTyrSerAlaAsp 40
 Db 90 CTCCCTCTCTGCCACCTCTCTGGGGTGGCTCAGCTCCACCGACGCTACTCTCTGAT 149
 QY 41 GlyLVsGlyThrGInSerHisProProGInuAspSerTTPLeuLVsPheArgSerGInuAsn 60
 Db 150 GCGAAGGCGACGTAGAGACACCGCCAGAGACAGCTGCTCAAGTTCCAGAGTGAGAAC 209
 QY 61 AenCySPhLeuTyrGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnPheVal 80
 Db 210 AACTGCTTCTGTATGGGGTCTTCAACGGCTATGATGGCACCGAGTGACCAACTTCGTG 269
 QY 81 AlAGInrArgLeuSerAlaGluLeuLeuLeuGInLeuAsnAlaGluHisAlaGluAla 100
 Db 270 GCCCAGCGGCGCTGCCACAGCTCTCGTGGGCCACGCTGAATGCCGACGCGCCAGGCC 329
 QY 101 AspValArgArgValLeuLeuGInAlaPheAspValValGluArgSerPheLeuGInSer 120
 Db 330 GATGTGCGCGCTGTGTCTGTCAGCGCTTCGATGTGTGAGAGAGAGCTTCCTGGAGTCC 389
 QY 121 IleAspAspAlaLeuAlaGluLVsAlaSerLeuGInSerGInLeuProGluGlyValPro 140
 Db 390 ATTGACGAGCGCTTGCTGAGAGGCGACGCTCCAGTCCGCAATTGCCAGAGGAGATCCCT 449
 QY 141 GInHisGInLeuProProGInTyrGInLVsIleLeuGInArgLeuLVsThrLeuGInArg 160
 Db 450 CAGCACCGACCTCCCTCTCAGTATCAGAAAGATCCTTGAGAGCTCAAGACGTTAGAGAGG 509
 QY 161 GluLVsSerGlyGlyAlaMetAlaValAlaValAlaLeuLeuAsnAsnLVsLeuTyrVal 180
 Db 510 GAAATTTTCGGAGGCGGCATGCGCTGTGGCGGCTCTTCACCAACAAGCTCTACGTC 569
 QY 181 AlaAsnValGlyThrAsnArgAlaLeuLeuCySlySerThrValAspGlyLeuGInVal 200
 Db 570 GCCAATGTGGTACAAACCGTGCCTTTATGCAAAATGCACAGTGGATGGGTGGAGGTG 629
 QY 201 ThrGInLeuAsnValAspHisThrThrGInuAsnGInuAspGluLeuPheArgLeuSerGIn 220
 Db 630 AACAGCTGAAGCTGGACCAACCAACAGAGAGAGAGATGAGCTTCCGCTTTCCGAG 689
 QY 221 LeuGlyLeuAspAlaGlyLVsIleLVsGInValGlyIleIleCySGLYGInGInSerThr 240
 Db 690 CTGGGCTTGATGTGGAAGATCAAGACAGGTGGGATCATCTGTGGGCGAGAGGACGCC 749

QY 241 ArgArgIleGlyAspTyrLVsValLVsTyrGlyTyrThrAspIleAspLeuLeuSerAla 260
 Db 750 CGCGGAGTCCGGGATTAACAAGTTAATATGCTACACGACGACATTGACTTCCAGCGCT 809
 QY 261 AlAlLVsSerLVsProIleIleAlaGluProGluLVsIleGlyAlaGInProLeuAspGly 280
 Db 810 GCCAAGTCCAAACCAATCATGTGCAGAGCCAGAAATCCATGGGGCACAGCGGTGGATGGG 869
 QY 281 ValThrGlyPheLeuValLeuMetSerGluGlyLeuTyrLVsAlaLeuGInAlaAlaHis 300
 Db 870 GTGACGGGCTTCTTGCTGCTGATGTGCGAGGGGTGTGACAAAGCCCTAGAGGAGCCCAT 929
 QY 301 GlyProGlyGInAlaAsnGInGluIleAlaAlaMetIleAspThrGluPheAlaLVsGIn 320
 Db 930 GGGGCTGGCAGGCGCAACAGAGATTGCTGCATGATTGACACTGATTTGCCAAGCAG 989
 QY 321 ThrSerLeuAspAlaValAlaGInAlaValValAspArgValLVsArgIleHisSerAsp 340
 Db 990 ACTTCCTTGAGCGCAGTGGCCAGGCCGTGTGACCGGGTGAAGCGCATCCACACCGAC 1049
 QY 341 ThrPheAlaSerGlyGlyAlaArgAlaArgPheCysProArgHisGluAspMetThrLeu 360
 Db 1050 ACCTTGCCAGTGTGTGGAGCGCTGCCAGGTTCTGCCCGCCGACAGAGACATGACCTTG 1109
 QY 361 LeuValArgAsnPheGlyTyrProLeuGlyGluMetSerGInProThrProSerProAla 380
 Db 1110 CTAGTGAGAACTTTGGCTACCCGCTGGGGCAATAGACACGCCACACGAGCCAGCC 1169
 QY 381 ProAlaIleGlyArgValTyrProValSerValProTyrSerSerAlaGInSerThr 400
 Db 1170 CCACTCCAGAGAGACGATGTACCTGTGTCTGTGCATACCTCCAGCCGCCAGACACC 1229
 QY 401 SerLVsThrSerValThrLeuSerLeuValMetProSerGInGlyGInMetValLVsGly 420
 Db 1230 AGCAAGACCGAGGACCTCTCCCTTGTATGCTCTCCAGGGCCAGATGTGCACGGG 1289
 QY 421 AlaHisSerAlaSerThrLeuAspGluAlaThrProThrLeuThrAsnGInSerProThr 440
 Db 1290 GCTCACAGTCTTCCACCCCTGGAGCAAGCCACCCCACTCCCAACCAAGCCCGACC 1349
 QY 441 LeuThrLeuGInSerThrAsnThrHisThrGInSerSerSerSerSerSerSerSer 460
 Db 1350 TTAACTTCGACGTCCACCAACACGACGACGAGAGAGAGCTTCCAGCTTGGAGAGGC 1409
 QY 461 LeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGlyArgValGluPro 480
 Db 1410 CTCTTCGCTCCGCGCCGCGCCCACTCGCTCCGCGCGGAGAGAGCGTGTGTAGGCC 1469
 QY 481 TyrValAspPheAlaGluPheTyrArgLeuTyrPheSerValAspHisGlyGluGInSerVal 500
 Db 1470 TATGTGACTTGTCTGATTTTAAACCGCTGTGAGCGTGAACCATGCGAGACAGAGCTGT 1529
 QY 501 ValThrAlaPro 504
 Db 1530 GTGACAGCACCG 1541
 RESULT 6
 HSU49928 3096 bp mRNA linear PRI 06-APR-1998
 LOCUS HSU49928
 DEFINITION Homo sapiens TAK1 binding protein (TAK1) mRNA, complete cds.
 ACCESSION U49928
 VERSION U49928.1 GI:1401125
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 3096)
 Shibuya,H., Yamaguchi,K., Shirakabe,K., Tonegawa,A., Gotoh,Y.,
 Ueno,N., Irie,K., Nishida,E. and Matsumoto,K.
 TAB1: an activator of the TAK1 MAPKK in TGF-beta signal
 transduction
 JOURNAL
 Science 272 (5265), 1179-1182 (1996)

```

MEDLINE 96216294
PUBMED 8638164
REFERENCE 2 (bases 1 to 3096)
TITLE Shibuya,H.
AUTHORS Direct Submission
JOURNAL Submitted (25-FEB-1996) Hiroshi Shibuya, Faculty of Pharmaceutical
Sciences, Hokkaido University, Nishi 6-chome, Kita 12, Kita-ku,
Sapporo, Hokkaido 060, Japan
FEATURES
source
1..3096
/organism="Homo sapiens"
/db_xref="taxon:9606"
gene
1..3096
/genes="TAB1"
CDS
21..1535
/genes="TAB1"
/note="activator for TAK1"
/codon_start=1
/product="TAK1 binding protein"
/protein_id="AAC12660.1"
/db_xref="GI:1401126"
/translation="MAAQRSLLSQEQPSWTDLDPLCHLSGVGSASNRYSADCKGT
ESHPPDSWLKPESENCEFLYGVNGYDGNRVTFNFAQRLSAELLGLQNAEHAADV
RRVLLQAFDWVERSFLESIDDLAEKASLSQSPGVPOHQPPQIKLERLKTLE
EISGGAMAVVALLNNKLXVANVTNRALLCKSTVDGLVQLNDVHTTENEDELFR
SQGLDAGKIKQVGIICGQESTRRIGDYKVKGYTDIDLLSAKSKPIIAEPEIHGAQ
PLDGVTFGLVMSSEGLIKALEAAHGPQAGIAAMIDLEFAKQTSLOAAOVVDV
KRTHSDTEASGGRRARFCPRHEDMTLLVRNFCYPLGEMSQPTSPAPAGGRRVPSV
PYSSAOSTSKTSVTLSPVPSQGMVNGHSASTLDEATPLITNSPILTLQSTNHT
QSSSSSDGKFRFRAHSLPPDGRVPEYVDFAEFYRLMWSVDHGEQSVVTA"
BASE COUNT 642 a 936 c 952 g 566 t
ORIGIN

Alignment Scores:
Pred. No.: 9,49e-151 Length: 3096
Score: 2580.00 Matches: 504
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-830-144-4 (1-504) x HSU49928 (1-3096)

Qy 1 MetAlaGlnArgSerLeuLeuGlnSerGluGlnProSerTrpThrAspAsp 20
Db 21 ATGCGCGCGCAGAGGAGAGCTTCTCAGAGTGAGCAGCAGCCAGCTGCAGATGAC 80
Qy 21 LeuProLeuCysHisLeuSerGlyValGlySerAlaSerAsnArgSerTyrSerAlaAsp 40
Db 81 CTGCCTCTCTGCCACCTCTCTGGGGTTGGCTCAGCCTCCAAACCGCAGCTACTCTGCTGAT 140
Qy 41 GlyLysGlyThrGluSerHisProProGluAspSerTrpLeuLysPheArgSerGluAsn 60
Db 141 GGC AAGGCGCACTGAGACCCACCCCGCAGAGACAGCTGGCTCAAGTTCAGAGTGAGAAC 200
Qy 61 AsnCysPheLeuTyrGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnPheVal 80
Db 201 AACTGCTTCTGTATGGGCTTTCAACGGCTATGATGGCAACCGAGTGACCAACTTCGTG 260
Qy 81 AlaGlnArgLeuSerAlaGluLeuLeuLeuGlyGlnLeuAsnAlaGluHisAlaGluAla 100
Db 261 GCCCAGCGGCTGTCGCGAGAGCTCCTCTGGGCGAGCTGAATGCCGAGCACGCCGAGGCC 320
Qy 101 AspValArgValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGluSer 120
Db 321 GATGTGGGGTGTCGTCTGAGGCCCTTCGATGTGGTGAGAGGAGCTTCCTGGAGTCC 380
Qy 121 IleAspAspAlaLeuAlaGluLysAlaSerLeuGlnSerGlnLeuProGluGlyValPro 140
Db 381 ATTGACACGCTTGGCTGAGAGGCAAGCCTCCAGTCGCAATTGCCAGNGGAGTCCCT 440
Qy 141 GlnHisGlnLeuProProGlnTyrGlnLysIleLeuGluArgLeuLysThrLeuGluArg 160

```

```

Db 441 CAGCACCACTGCCTCTCTCAGTATCAGAAAGATCCTTGCAGACTCAAGACGCTTAGAGAGG 500
Qy 161 GluIleSerGlyGlyAlaMetAlaValValAlaValLeuLeuAsnAsnLysLeuTyrVal 180
Db 501 GAAATTTTCGGAGGGGCCATGGCGTTGTGGGGTCTTCTTCAACAACAAGCTCTACGTC 560
Qy 181 AlaAsnValGlyThrAsnArgAlaLeuLeuLeuCysLysSerThrValAspGlyLeuGlnVal 200
Db 561 GCCAATGTCTGGTACAAACCGTGCACCTTTTATGCAAAATCGACAGTGGATGGGTTCAGGTG 620
Qy 201 ThrGlnLeuAsnValAspHisThrThrGluAsnGluAspGluLeuPheArgLeuSerGln 220
Db 621 ACACAGCTGAACGTGGACCAACACAGAGAACGAGGATGAGCTCTTCCGCTTTTCGCAG 680
Qy 221 LeuGlyLeuAspAlaGlyLysIleLysGlnValGlyIleIleCysGlyGlnGluSerThr 240
Db 681 CTGGGCTTTGGATGCTGGAAAGATCAAGCAGGTGGGATCATCTGTGGCGCAGAGAGCAC 740
Qy 241 ArgArgIleGlyAspTyrLysValLysTyrGlyTyrThrAspIleAspLeuLeuSerAla 260
Db 741 CGCGGATCGGGGATTACAAGGTTAAATATGGCTACACGGACATTGACCTTCTCAGCGCT 800
Qy 261 AlaLysSerLysProIleIleAlaGluProGluIleHisGlyAlaGlnProLeuAspGly 280
Db 801 GCCAAGTCCAAACCAATCATCGCAGAGCCAGAAATCCATGGGGCACAGCCGCTGGATGG 860
Qy 281 ValThrGlyPheLeuValLeuMetSerGluGlyLeuTyrLysAlaLeuGluAlaAlaHis 300
Db 861 GTGACGGCTCTTCTGGTCTCATGTGCGAGGGGTTGTACAAGGCCCTTAGAGGACGCCAT 920
Qy 301 GlyProGlyGlnAlaAsnGlnGluIleAlaAlaMetIleAspThrGluPheAlaLysGln 320
Db 921 GGGCTCGGGCAGGCCAACCCAGGAGATTGCTCGGATGATTGACACTGAGTTTGCCACAGCAG 980
Qy 321 ThrSerLeuAspAlaValAlaGlnAlaValValAspArgValLysArgIleHisSerAsp 340
Db 981 ACCTCCTCGCAGCGCAGTGGCCAGCCGCTCGTGACCCGGTGAAGCGCATCCACAGCGCAC 1040
Qy 341 ThrPheAlaSerGlyGlyGluArgAlaArgPheCysProArgHisGluAspMetThrLeu 360
Db 1041 ACCTTGGCAGTGTGGGGGCGTGCCAGGTTCTGCCCGCGCAGGAGGATGACCCCTG 1100
Qy 361 LeuValArgAsnPheGlyTyrProLeuGlyGluMetSerGlnProThrProSerProAla 380
Db 1101 CTAGTGAGGAACTTTGGCTACCCGCTGGCGCAATGAGCCAGCCACACCCGAGCCGACCC 1160
Qy 381 ProAlaAlaGlyArgValTyrProValSerValProTyrSerSerAlaGlnSerThr 400
Db 1161 CCAGCTGCAGGAGGACGAGTGTACCTGTGTGTGTCATACTCCAGCGCCCGCAGAGCAC 1220
Qy 401 SerLysThrSerValThrLeuSerLeuValMetProSerGlnGlyGlnMetValAsnGly 420
Db 1221 AGCAAGACACGCTGACCCTCTCCCTTGTATGCCCTCCAGGGCCAGATGGTCAACGGG 1280
Qy 421 AlaHisSerAlaSerThrLeuAspGluAlaThrProThrLeuThrAsnGlnSerProThr 440
Db 1281 GCTCAGCTGCTTCACCCCTCGACGAGCCACCCGCCACCTCACCAACCAAGCCCGACC 1340
Qy 441 LeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSerAspGlyGly 460
Db 1341 TTAAACCTTCGAGTCCCAACACGACGACGAGCAGCAGCTCCAGCTCTGACGGAGGC 1400
Qy 461 LeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGlyArgValGluPro 480
Db 1401 CTCCTCCGCTCCCGGCCCGCCACTCGCTCCCGCTCGCGAGGACGGTCTGTGTAGGCC 1460
Qy 481 TyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGlyGluGlnSerVal 500
Db 1461 TATGTGGACTTTGTGTAGTTTTACCGCTCTTGGAGCGTGGACCATGGCGACGAGCGGTG 1520
Qy 501 ValThrAlaPro 504
Db 1521 GTGACAGCACCG 1532

```

RESULT 7									
LOCUS	AR058302		1560 bp	DNA				linear	PAT 29-SEP-1999
DEFINITION	Sequence 5 from patent US 5837819.								
ACCESSION	AR058302								
VERSION	AR058302.1		GI:5983879						
KEYWORDS									
SOURCE									
ORGANISM	Unknown.								
REFERENCE	Unclassified.								
AUTHORS	1 (bases 1 to 1560)								
TITLE	Matsumoto,K. and Nishida,E.								
JOURNAL	TAB1 protein								
FEATURES	Patent: US 5837819-A 5 17-NOV-1998;								
SOURCE	Location/Qualifiers								
	1..1560								
	/organism="unknown"								
BASE COUNT	333 a	468 c	480 g	279 t					
ORIGIN									

Alignment Scores:

Pred. No.:	8,14e-151	Length:	1560
Score:	2575.00	Matches:	503
Percent Similarity:	99.80%	Conservative:	0
Best local Similarity:	99.80%	Mismatches:	1
Query Match:	99.81%	Indels:	0
DB:	6	Gaps:	0

US-09-830-144-4 (1-504) X AR058302 (1-1560)

OY	1	MeAlaAlaGlnARGArgSerLeuLeuGlnSerGluGlnGlnProSerTPMThrAspAsp	20
Db	30	ATGGGGGGCGAGAGAGAGAGCTTGCTGCAGGTGACGACGACGACCTGGACGATGAC	89
OY	21	LeuProLeuCySHisLeuSerGlyValGlySerAlaSerAsnArgSerTyrSerAlaAsp	40
Db	90	CTGGCTCTTGCGACCTCTCTGGGGTGGCTCAGCTTCCAAACCCGACCTACTCTGGTAT	149
OY	41	GlyIysGlyThrGlnSerHisProProGlnAspSerPheLeuLysPheArgSerGluAsn	60
Db	150	GGCAAGGGACACTGAGAGCGACCCCGCAGAGAGACGATGCTCAAGATTCAAGAGTGAAC	209
OY	61	AsnCySPheLeuTyrGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnPheVal	80
Db	210	AACTGCTTCCGTATGGGGTCTTCAACGGCTATGATGGCAACCGAGTGAACCACTTGGTG	269
OY	81	AlaGlnArgLeuSerAlaGluLeuLeuLeuGlyGlnLeuAsnAlaGluHisAlaGluAla	100
Db	270	GCCCGAGGGGCTGTCCGACAGACTCTGCTGGGCCACGCTGAATGCGGAGACGCGGAGGCC	329
OY	101	AspValArgArgValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGlnSer	120
Db	330	GATGTGGCGGCGTGTCTGCTGAGGCTTTCGATGGTGGGAGAGAGACTTCTCGAATCC	389
OY	121	IleAspAspAlaLeuAlaGluLysAlaSerLeuGlnSerGlnLeuProGluGlyValPro	140
Db	390	ATTGACGACCGCTTGGCTGAGAGAGCAAGCTTCAAGTCGCAATGGCCAGAGGAGTCTCT	449
OY	141	GlnHisGlnLeuProProGlnTyrGlnLysIleLeuGluArgLeuLysThrLeuGluArg	160
Db	450	CAGCACACAGCTCCCTCTCAGTATAGAAAGTCCCTTGAGACACTCAAGAGCTTGAAGAGG	509
OY	161	GluIleSerGlyValAlaMetAlaValAlaValLeuLeuAsnAsnLysLeuTyrVal	180
Db	510	GAATTTTGGGGGGGGCCATGGCCGTTGTGGCGGCTTTCCAACAACAAGCTCTAGTGC	569
OY	181	AlaAsnValGlyThrAsnArgAlaLeuLeuCyLysSerThrValAspGlyLeuGlnVal	200
Db	570	GCCATGTGCGTACAAACCGTGCACCTTTATGCAAAATGACAGTGGATGGGTTCAGAGTG	629
OY	201	ThrGlnLeuAsnValAspHisThrThrGluAsnGluAspGluLeuPheArgLeuSerGln	220

Db	630	ACACAGCTGAACGTGGAGACCAACACAGAGAGACAGATGAGCTCTTCCGCTTTCGAG	689
QY	221	LeuGlyLeuAspAlaGlyLysIleGlnValGlyIleIleGlyGlnGlnSerThr	240
Db	690	CTGGGCTTGATGTGCTGGAAGAATCAAGACGCTGGGATCATCTGTGGGACAGAGCAC	749
QY	241	ArgArgIleGlyAspTyrLysValLysTyrGlyTyrThrAspIleAspLeuSerIa	260
Db	750	CGGCGGATCGGGGATTCACAGAGTTAAATATGGCTACACAGGACATTGACTTCTCAGCGCT	809
QY	261	AlaLysSerLysProIleIleAlaGlnProGlnIleHisGlyAlaGlnProLeuAspGly	280
Db	810	GCCAGTGTCAACCATCATCATCGACAGGCAAAATCCATGGGCGACAGCCCTGATGG	869
QY	281	ValThrGlyPheLeuValIleuMetSerGlnGlyLeuTyrLysAlaLeuGlnAlaHis	300
Db	870	GTCAGCGGCTCTTGATGCTGATGTCGAGGGGTTGTACAGGCGCTTGAGGAGCCCAT	929
QY	301	GlyProGlyGlnAlaAsnGlnIleIleAlaMetIleAspThrGlnPheAlaGln	320
Db	930	GGGCTGGGAGGCGCACAGAGAAATGCTGCGATGATGACACTGACTTTGGCCAGAG	989
QY	321	ThrSerLeuAspAlaValAlaGlnAlaValAspArgValLysArgIleHisSerAsp	340
Db	990	ACCTCCCTGGAGACGATGGCCCAAGCCCTCGTGCAGCGGGTGAAGCGATCCACAGCGAC	1049
QY	341	ThrPheAlaSerGlyGlyGlnArgAlaArgPheCysProArgHisGlnAspMetThrLeu	360
Db	1050	ACCTTCGGCAGTGGGGGAGCGTCCGAGTTGTGCCCCGGGACGAGACATGACCTTG	1109
QY	361	LeuValArgAsnPheGlyTyrProLeuGlnGlnMetSerGlnProThrProSerProIa	380
Db	1110	CTAGAGAGAACTTGGCTACCCGCTGGGCGAAATGAGCCAGCCACACGAGCCAGACC	1169
QY	381	ProAlaAlaGlyGlyArgValTyrProValSerValProTyrSerSerAlaGlnSerThr	400
Db	1170	CCAGCTGCAGAGAGACGAGTACCTCTGTCTGTGCTCATCTCCAGGCCCCACAGCAC	1229
QY	401	SerLysThrSerValThrLeuSerLeuValMetProSerGlnGlyMetValAsnGly	420
Db	1230	AGCAAGACAGAGTGACCTCTCTCTGTATGCCCTCCAGGGCCAGATGATCAACGGG	1289
QY	421	AlaHisSerAlaSerThrLeuAspGlnAlaThrProThrLeuThrAsnGlnSerProThr	440
Db	1290	GCTCACTGCTCTCCACCTTGACGAAACCAACCCCACTTCACCAACCAAGCCGAC	1349
QY	441	LeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSerSerAspGlyGly	460
Db	1350	TTTAACTTCGTCATCCACCAACGACGACGACGACGACGACGACGACGACGACGAC	1409
QY	461	LeuPheArgSerArgProAlaHisSerLeuProProGlyGlnAspGlyArgValGlnPro	480
Db	1410	CTCTTCGCTCCGGCCCGCCCACTCGCTCCGCTGGCGAGAGAGCGTGTGTAGCCC	1469
QY	481	TyrValAspPheAlaGlnPheTyrArgLeuThrPheSerValAspHisGlyGlnGlnSerVal	500
Db	1470	TATGTGACATTGTGCGATTTTACCGCTCTGGAGCGTGAGACCATGGGAGACAGAGCTG	1529
QY	501	ValThrAlaPro 504	
Db	1530	GTGACAGCACCG 1541	
RESULT 8			
LOCUS	AR088276	1560 bp	DNA linear PAT 07-SEP-2000
DEFINITION	Sequence 5 from patent US 5989862.		
ACCESSION	AR088276		
VERSION	AR088276.1 GI:10015039		
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 1560)		

AUTHORS Matsumoto,K. and Nishida,E.
TITLE Tab1 protein and DNA coding therefor
JOURNAL Patent: US 5989862-A 5 23-NOV-1999;
FEATURES Location/Qualifiers
source 1..1560
/organism="unknown"

BASE COUNT 333 a 468 c 480 g 279 t
ORIGIN

Alignment Scores:
Pred. No.: 8.14e-151 Length: 1560
Score: 2575.00 Matches: 503
Percent Similarity: 99.80% Conservative: 0
Best Local Similarity: 99.80% Mismatches: 1
Query Match: 99.81% Indels: 0
DB: Gaps: 0

US-09-830-144-4 (1-504) x AR088276 (1-1560)

Qy 1 MetAlaAlaGlnArgSerLeuLeuGlnSerGluGlnProSerTrpThrAspAsp 20
Db 30 ATGGCGCGCAGAGGAGGAGCTTGTCTGACAGGTGAGCAGCAAGCTGGACAGATGAC 89
Qy 21 LeuProLeuCysHisLeuSerGlyValGlySerAlaSerAsnArgSerTyrSerAlaAsp 40
Db 90 CTGCCTCTCTGCCACCTCTCTGGGGTGGCTCAGCCTCCAACCGCAGCTACTCTGCTGAT 149
Qy 41 GlyLysGlyThrGluSerHisProGluAspSerTrpLeuLysPheArgSerGluAsn 60
Db 150 GGCAAGGGCACTGAGAGCCACCGCCAGAGGACAGATGGCTCAAGTTCAAGAGTGAGAAC 209
Qy 61 AsnCysPheLeuTyrGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnPheVal 80
Db 210 AACTGCTCTCTATGGGCTTTCAACGGCTATGATGGCAACCGAGTGACCAACTTCGTG 269
Qy 81 AlaGlnArgLeuSerAlaGluLeuLeuLeuGlyGlnLeuAsnNlaGluHisAlaGluAla 100
Db 270 GCCCAGCGGCTGTCGAGAGCTCTCTGGCCAGCTGAATGCCGAGCACGCCGAGGCC 329
Qy 101 AspValArgArgValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGluSer 120
Db 330 GATGTCGGCGCTGTCTGTCTGAGGCTTCATGTGGTGAGAGGAGCTTCCTGGAGTCC 389
Qy 121 IleAspAspAlaLeuAlaGluLysAlaSerLeuGlnSerGlnLeuProGluGlyValPro 140
Db 390 ATTGACAGCCCTTGGCTGAGAAGGCAAGCCTCCAGTCGCAATTGCCAGAGGGAGTCCCT 449
Qy 141 GlnHisGlnLeuProProGlnTyrGlnLysIleLeuGluArgLeuLysThrLeuGluArg 160
Db 450 CAGCACAGCTGCCTCTCTCAGTATCAGAGATCCTTTGAGAGACTCAAGACGTTAGAGAGG 509
Qy 161 GluIleSerGlyGlyAlaMetAlaValAlaValAlaLeuLeuAsnAsnLysLeuTyrVal 180
Db 510 GAAATTCGGAGGGCCATGGCCGTTGTGGCGGCTCTTCTCAACAACAAGCTCTACAGTC 569
Qy 181 AlaAsnValGlyThrAsnArgAlaLeuLeuCysLysSerThrValAspGlyLeuGlnVal 200
Db 570 GCCAATGTCGTACAAACCGTCACATTTATGCAATCGACAGTGGATGGGTTCAGAGTG 629
Qy 201 ThrGlnLeuAsnValAspHisThrThrGluAsnGluAspGluLeuPheArgLeuSerGln 220
Db 630 ACACAGCTGAACCTGGACCAACACACAGAGAACAGGATGAGCTCTTCGCTCTTTCGCAG 689
Qy 221 LeuGlyLeuAspAlaGlyLysIleLysGlnValGlyIleCysGlyGlnGluSerThr 240
Db 690 CTGGGCTTGATCTGGAAGATCAAGCAGAGTGGGATCATCTGTGGCGAGGAGACACC 749
Qy 241 ArgArgIleGlyAspTyrLysValLysTyrGlyTyrThrAspIleAspLeuLeuSerAla 260
Db 750 CGCGGATCGGGATTACAAGGTTAAATATGGCTACACGGACATTGACCTTCTCAGCGT 809
Qy 261 AlaLysSerLysProIleLeuAlaGluProGluIleHisGlyAlaGlnProLeuAspGly 280

Db 810 GCCAAGTCCAAACCAATCATCGCAGAGCCAGAAATCCATCGGGGCACAGCCGCTGGATGGG 869
Qy 281 ValThrGlyPheLeuValLeuMetSerGluGlyLeuTyrLysAlaLeuGluAlaAlaHis 300
Db 870 GTGACGGCTTCTTGGTGCTGATGTCGAGGGGTTGTACAAGGCCCTTAGAGGCAGCCCAT 929
Qy 301 GlyProGlyGlnAlaAsnGlnGluIleAlaAlaMetIleAspThrGluPheAlaLysGln 320
Db 930 GGGCTTGGCAGGCGCAACAGAGATTGCTGCGATGATTGACATGAGTTTGCCACAGCAG 989
Qy 331 ThrSerLeuAspAlaValAlaGlnAlaValValAspArgValLysArgIleHisSerAsp 340
Db 990 ACCTCCCTGGACCGCAGTGCCCGACCGCGTCTGTGACCGGGTGAAGCGCATCCACAGCGAC 1049
Qy 341 ThrPheAlaSerGlyGlyGluArgAlaArgPheCysProArgHisGluAspMetThrLeu 360
Db 1050 ACCTTCGCCAGTGTGGGGAGCGTCCAGGTTCTGCCCCCGGCACAGGACATGACCCCTG 1109
Qy 361 LeuValArgAsnPheGlyTyrProLeuGlyGluMetSerGlnProThrProSerProAla 380
Db 1110 CTAGTGAGGAACTTTGGCTACCCGCTGGCGAAATGAGCCAGGCCACACCGAGCCAGGCC 1169
Qy 381 ProAlaAlaGlyGlyArgValTyrProValSerValProTyrSerSerAlaGlnSerThr 400
Db 1170 CCAGTCGAGGAGGACGAGTGATCCCTGTGTGTGTGCATCTCCAGCGCCCGCAGAGCACC 1229
Qy 401 SerLysThrSerValThrLeuSerLeuValMetProSerGlnGlyGlnMetValAsnGly 420
Db 1230 AGCAAGACCAGCGTGACCTCTCCCTTGTCTATGCTCCAGGSCCAGATGTTCAACGGG 1289
Qy 421 AlaHisSerAlaSerThrLeuAspGluAlaThrProThrLeuThrAsnGlnSerProThr 440
Db 1290 GCTCACAGTGCTTCCACCTTGACGAAAGCCACCCGCCCTCCACCAACCAAGCCCGACC 1349
Qy 441 LeuThrLeuGlnSerThrAsnThrHisThrClnSerSerSerSerSerSerAspGlyGly 460
Db 1350 TTAACCTTGCACTCCACCAACACACACACGACGACGAGCAGCTCCAGCTCTGACGAGGC 1409
Qy 461 LeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGlyArgValGluPro 480
Db 1410 CTCCTTCGCTCCCGCGCGCCCACTCGCTCCCGCTCGCAGGACGGTGTGTTGAGCCC 1469
Qy 481 TyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGlyGluGlnSerVal 500
Db 1470 TATGTGAGACTTTGCTGAGTTTACCGCTCTGAGCGCTGACCATGCGCAGCAGAGCGTG 1529
Qy 501 ValThrAlaPro 504
Db 1530 GTGACAGCACCG 1541

RESULT 9
LOCUS AR116884 1560 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 5 from patent US 6140042.
ACCESSION AR116884
VERSION AR116884.1 GI:14097790
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1560)
AUTHORS Matsumoto,K. and Nishida,E.
TITLE TAB1 protein and DNA coding therefor
JOURNAL Patent: US 6140042-A 5 31-OCT-2000;
FEATURES Location/Qualifiers
source 1..1560
/organism="unknown"
BASE COUNT 333 a 468 c 480 g 279 t
ORIGIN
Alignment Scores:
Pred. No.: 8.14e-151 Length: 1560
Score: 2575.00 Matches: 503

Db 1495 GCACCT 1500

RESULT 11
LOCUS XLU92031 1926 bp mRNA linear VRT 18-APR-1998
DEFINITION Xenopus laevis TAK1 binding protein TAB1 mRNA, complete cds.
ACCESSION U92031
VERSION U92031.1 GI:3057037
KEYWORDS
SOURCE Xenopus laevis.
ORGANISM Xenopus laevis.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus.
REFERENCE
1 (bases 1 to 1926)
Shibuya,H., Iwata,H., Masuyama,N., Gotoh,Y., Yamaguchi,K., Irie,K.,
Matsumoto,K., Nishida,E. and Ueno,N.
Role of TAK1 and TAB1 in BMP signaling in early Xenopus development
EMBO J. 17 (4), 1019-1028 (1998)
JOURNAL
MEDLINE 98130593
PubMed 9463380
REFERENCE
2 (bases 1 to 1926)
Shibuya,H.
Direct Submission
JOURNAL Submitted (05-MAR-1997) Faculty of Sciences, Hokkaido University,
Nishi 6-chome, Kita 12, Kita-ku, Sapporo, Hokkaido 060, Japan
FEATURES
source
1..1926
Location/Qualifiers
/organism="Xenopus laevis"
/db_xref="taxon:8355"
38..1534
/note="TAK1 binding protein"
/codon_start=1
/product="TAB1"
/protein_id="AAC14009.1"
/db_xref="GI:3057038"
translation="MAAPRRNLHSQSWTDLPICNLISGVGASNOTYNSRGLKDEH
PYEDNWKIRGDNNTLYLVGRNXYEGTATSEVGRRLAEELLGLDPTDADAVHKY
LQADVVERFLESIDDLAKSLSQLEPGTLHQLPQYQKIVRLNLEKEYI
GGAMVIVLVNSKLIVANVNGTRALILCKSTVDGIVQLVADHTENEDILLSOL
GLDTTKIKOVGIVGGQSTRIGDKYKVNFDLELSTASKXPITAEPEHGCQPLD
GVDTGLVMSSEGLYKALPSAHPGOANOELAMLTETPAKGVSDENVAOLVEEVKSI
HDDTPASGERAKYCSKHEMDTLVRNIGYPLQETSPATGPGGRLPYSPVSSA
ONTSKTSVTLISLWPSQGPVNVGNTSSSTLDGTTSTLOSPEATLOSTRTTQSSSSS
DGGFRSRPLPLQPEDGKVEPVDFTFYRLWMAERNDPTLTAAQ"
BASE COUNT 556 a 462 c 456 g 451 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 2,14e-110 Length: 1926
Score: 1924.00 Matches: 378
Percent Similarity: 85.71% Conservative: 54
Best Local Similarity: 75.00% Mismatches: 64
Query Match: 74.57% Indels: 8
DB: Gaps: 4

US-09-830-144-4 (1-504) x XLU92031 (1-1926)

Qy 1 MetAlaAlaGlnArgSerLeuLeuGlnGlnInProSerTrpThrAsp 20
Db 38 ATGGCTGGCCCGAAGAGATCTTCTGCATAGTCAG-----AGCTGCAGCATGAT 88

Qy 21 LeuProLeuCyHisIleuSerGlyValGlySerAlaSerAsnArgSerTrpSerAlaAsp 40
Db 89 CTCCTCTCTGCACCTTCTCGGGGTTGGATCTGCTTCCATCAGACTTACCACTCAGAG 148

Qy 41 GlyLysGlyThrGluSerHisProProGluAspSerTrpLeuLysPheArgSerGluAsn 60
Db 149 GGGCTTTGGGAAAGACGACCATTAAGAATACTGATTAAGTTCAGGGGCGACAT 208

Qy 61 AsnCySPheLeuTyGlyValPheAsnGlyTyAspGlyAsnArgValThrAsnPheVal 80
Db 209 AATATTATCTTTATGAGAGTGTTTAAACNGTATGAGGACCAAGCCACGCTTTGTG 268

Qy 81 AlaGlnArgLeuSerAlaGluLeuLeuGlyGlnIleuAsnAlaGluHisAlaGluAla 100
Db 269 GGTCAAGCGGCTTGACGACGAGAACTACTTCTGGCGAGTTGGACCTTGACAGATGCA 328

Qy 101 AspValArgArgValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGluSer 120
Db 329 GAGGTCCAAAGATATTGCTACAGGCACTTGATGTTGTTAAAGAAAGCTTCTGGAGTCC 388

Qy 121 IleAspAspAlaLeuAlaGluLysAlaSerLeuGlnSerGlnIleuProGluGlyValPro 140
Db 389 ATCGATGATGCTTACGAGAAAAGCCAGCCTTGCTGCCAATCTTCAGAGGCTTACTCTG 448

Qy 141 GlnHisGlnIleuProProGlnTyTrpGlnLysIleLeuGluArgLeuLysThrLeuGluArg 160
Db 449 CACCAACACCTTCCAAACCCAGTCCAAAGAAATGATGTGACCGGCTTAAACATCTTGAGAG 508

Qy 161 GluIleSerGlyGlyAlaMetAlaValAlaValLeuLeuAsnAsnLysLeuTyTrpVal 180
Db 509 GAAATATATGGGGAGGCGACATGCTATTTGTGTTGTAACAGCAACTTATATGT 568

Qy 181 AlaAsnValGlyThrAsnArgAlaLeuLeuCyLysSerThrValAspGlyLeuGlnVal 200
Db 569 GCAATGTCCGAAACAACAGACGACTTTGTGTAATCACTTACATGCGCTTCAAGT 628

Qy 201 ThrGlnLeuAsnValAspHisThrThrGluAsnGluAspGluLeuPheArgLeuSerGln 220
Db 629 ACTCAGCTCAATGCTGACACACACACGAAATGAAGATAGATATGTCGTCTGTCAA 688

Qy 221 LeuGlyLeuAspAlaGlyLysIleLysGlnValGlyIleLeuGlyGlnGluSerThr 240
Db 689 TTGGGCTGGACACTACAAAGATTTAAACAGGTGGGGTTTGTGAGGCCGCAACGACT 748

Qy 241 ArgArgIleGlyAspTyTrpLysValLysTyGlyTyThrAspIleAspLeuLeuSerAla 260
Db 749 CGCAGATTGGAGACTACAAAGCTCAAGGTCAAGTCAACTCAATGACCTTACGCTCAACACA 808

Qy 261 AlaLysSerLysProIleIleAlaGluProGlnIleHisGlyAlaGlnProLeuAspGly 280
Db 809 GCCAAGTCCAAAGCTATTTACTGCAAGCCAGAGATCCATGCTGCCAGCATTAAGCGGT 868

Qy 281 ValThrGlyPheLeuValLeuMetSerGlnGlyLeuTyTrpValAlaGlnIleAlaHis 300
Db 869 GTCCAGAGATTTTGGTCTTAATGCTGAAAGACTTTCAGAGCTTCGAGTACGCCAT 928

Qy 301 GlyProGlyGlnAlaAsnGlnIleAlaAlaMetIleAspThrGluPheAlaLysGln 320
Db 929 GGGCTGACAGGCGCAACGAGAAATAGCGCATGATTCGCCAGGTTTGGCAAGCAG 988

Qy 321 ThrSerLeuAspAlaValAlaGlnAlaValValAspArgValLysArgIleHisSerAsp 340
Db 989 GTTTCCTCTGATGTAAGTGGCAGAGGCTTGTGTGAGAGGGTTAAAGGATTCACCATGAT 1048

Qy 341 ThrPheAlaSerGlyGlyGluArgAlaArgPheCySPheArgHisGluLysPheMetThrLeu 360
Db 1049 ACGTTTCAAGTGTGGCGAAGCGGCAAAATATTGAGTAAAGCAGAAACATGACGCTA 1108

Qy 361 LeuValArgAsnPheGlyTyTrpProLeuGlyGluMetSerGlnIleProThrProSerProAla 380
Db 1109 CTGTGCGGAACCTGGGCTTACCGGCTTCAGAGATACAGCCCCCACCACACTCAG----- 1162

Qy 381 ProAlaAlaGlyArgValTyProValSerValProTySerSerAlaGlnIleSerThr 400
Db 1163 CCCACTCAAGGTGAGCCTTATATCCAGTGTCTGTGCCATATTCAGAGTCTCGAATACC 1222

Qy 401 SerTyThrSerValThrLeuSerLeuValMetProSerGlnGlyGluMetValAsnGly 420
Db 1223 AGTAAACAAAGTGCACACTGCTACGTGATGCCATCACAGGTCCTCAATGGGATGGA 1282

Qy 421 AlaHisSerAlaSerThrLeuAspGluAlaThrProThrLeuThrAsnGlnIleSerProThr 440
Db 1283 ACAACAGCAGCTCAACACTGGATGGAAACACAGCTCAACTG-----CAATCTCCAACT 1336

Gene

10049. 42131
/gene="d4407f17.1"
join(10049. 10084,25255. 25391,25749. 25902,27040. 27126
27960. 28098,28981. 29094,29768. 29879,32076. 32220,
36952. 37174,38270. 38432,40264. 42131)
/gene="d4407f17.1"
/product="d4407f17.1 (TAK1 (TAK1 binding protein 1))"
/note="match: cDNAs: Em:U9201 Em:U4928; match: ESTs:
Em:HA1299 Em:AA726609 Em:AA019355 Em:AA16567 Em:AL65435
Em:AA189149 Em:AI014848 Em:AA430502 Em:AA745822
Em:AA411895 Em:AA411894 Em:AA326751 Em:AA430621 Em:HS93737
Em:AA430462 Em:AA806521 Em:R40486 Em:AA019356 Em:AA157525
Em:AA028980 Em:U34667 Em:R83147 Em:AI69211 Em:AA806517
Em:AA147685 Em:AA48466 Em:Z43346 Em:AI656463 Em:W93352
Em:AA828095 Em:R60090 Em:AA280338 Em:AA679208 Em:AA442978
Em:R60391 Em:AI085576 Em:W95394 Em:R13055 Em:HA8607
Em:F02762"
/evidence=not experimental
join(10052. 10084,25255. 25391,25749. 25902,27040. 27126
27960. 28098,28981. 29094,29768. 29879,32076. 32220,
36952. 37174,38270. 38432,40264. 40471)
/gene="d4407f17.1"
/note="match: proteins: Tr:Q15750 Tr:Q73614"
/codon_start=1
/evidence=not experimental
/product="d4407f17.1 (TAK1 (TAK1 binding protein 1))"
/protein_id="CAB55304.1"
/db_xref="GI:5834565"
/translation="MAAQRRLSQEQPWTDDLPCLHSGVGSASRNSYSADCKGT
ESHPEPDSKLRSENNCPDLGVFNQYDGNRTNVAORLSMELLGOLNHEADV
RVLLOAFVVRSEFLPSIDILAEKASLOSOLPQYQHOHLPPQYKILRKLTER
ELSGMAAVAVALLNNKKLVYAVGTRRLCKSTGYDGLQVLDNPTNDEDELRL
SOLGDAKIKGVGICQESTRRIGDYKVKYKGTIDILLSAKSKPIIARPEHGAO
PLDGTGFLVMSBGLYKALEAHAGRGANDQIAIMIDTERFAKQTSLDVAQAVDRA
KRIHSDTFSKSGSRARPCRHEDMTLLVKNFQYPISESGSQPPPSAPAGRVVPSV
PYSASQTSKTSVTLTLWPSQGVQWNGHSASTIDENATPLTINQSPILTLQSTVTH
QSSSSSDGGLFRSRPAHSLPGEGDGRVPEYDFAEFLRMSVDHGQSVTAP"
10423. 10665
/note="MIR repeat: matches 3. 261 of consensus"
10894. 11003
/note="L2 repeat: matches 2483. 2587 of consensus"
11004. 11310
/note="AluY repeat: matches 2. 308 of consensus"
11311. 11418
/note="L2 repeat: matches 2587. 2710 of consensus"
12433. 12563
/note="FLAM_C repeat: matches 2. 133 of consensus"
12622. 13098
/gene="d4407f17.1"
/note="match: GSS: Em:AQ601543"
12823. 13100
/note="AluS9 repeat: matches 1. 279 of consensus"
13182. 13227
/note="MIR repeat: matches 86. 131 of consensus"
13351. 13887
/note="L2 repeat: matches 2087. 2749 of consensus"
13888. 14192
/note="AluS9 repeat: matches 1. 298 of consensus"
14193. 14515
/note="L2 repeat: matches 1790. 2087 of consensus"
14551. 14831
/note="AluJ repeat: matches 1. 298 of consensus"
14601. 16671
/note="AluS1 repeat: matches 35. 305 of consensus"
16869. 17181
/note="AluY repeat: matches 1. 310 of consensus"
17473. 17722
/note="L1 repeat: matches 4070. 4338 of consensus"
17724. 18033
/note="AluY repeat: matches 1. 135 of consensus"
18034. 18168
/note="AluY repeat: matches 1. 135 of consensus"

Alignment Scores:		3.32e-23	Length:	69660
Pred. No.:		547.50	Matches:	296
Score:		18.93%	Conservative:	44
Percent Similarity:		16.48%	Mismatches:	109
Best Local Similarity:		21.22%	Indels:	1348
Query Match:		9	Gaps:	19
DB:				
US-09-830-144-4 (1-504) x HS407E17 (1-69660)				
QY	8	LeuLeudlnserscIugInglInpRoserTrphraSPaspleuProleuCySHsleuser	27	
Db	25243	GTCTTTCCCGAGAGACAGCCAGCTGAAGATGAATGACTCTCTCTCTCCACTCTCTCT	25302	
QY	28	GlyValGlySerIaSerAsnArgserTySerIaaspGlyIysGlyThrIuserHis	47	
Db	25303	GGGGTTTGCGCTACACCTCCAGCCGACGACTACTCTGTATGTSCAAAGGSCATGAGACCCAC	25362	
QY	48	ProProGlnAspSerTrpLeuLysPhe-----	56	
Db	25363	CCGCGAGAGAGACAGCTGAGCTCAAGTTCAAGTGTGTGTGTCAGACATTTCTGTGTGGGCC	25422	
QY	56	-----	56	
Db	25423	GGGAGTTGGTTGGTTGTTGCCAAGCAAGAAACACCGACTTGACGCTTTCTGTATGG	25482	
QY	56	-----	56	
Db	25483	CTTGCCAGTACATGTGGCCCGCTGAGAGATGGCTCTGTCTGTCTTGTGCCAAGGCGCTG	25542	
QY	56	-----	56	
Db	25543	CTCTGATGGGGTAGCGTGAGCATGGGGAGAGAGGATATCCAGAATGTATAGCCAGAGTG	25602	
QY	56	-----	56	
Db	25603	AAATGATGGCTAAAGCAGGGGGAGCCAGAGAGGCCCTTGAAGCTGACGTCTGTGCTT	25662	
QY	56	-----	56	
Db	25663	TAGTCTCCCAATTCTTTCCTTCTCTCCACTCCGCTGAGACCTCGGTCTCAGGC	25722	
QY	57	-----ArgSerGlnAsnAsnCysPheLeuTyrglyValPhe	68	
Db	25723	CTCCCTCTGGCCCTCTCCCTCTTCCAGAGTAGAAGCAACTGCTTCTGTATGGGGCTTTC	25782	
QY	69	AsnGlyTyTrAspGlyAsnAsnArgValThrAsnHeValIaGlnArgLeuSerIaGlyLeu	88	
Db	25783	AACGGCTATGATGGCAACCGAATGACCAACTTCTGTGCCACGAGCGCTGCCGAGGCTC	25842	
QY	89	LeuLeuGlyGlnLeuAsnIaGlyHisIaGlyIaaspValArgArgValLeuLeuGln	108	
Db	25843	CTGCTGGGCCAGTGAATGCCAGACGCCAGGCGCAATGTGGGGCTGTGTCTGTCCAG	25902	
QY	108	-----	108	
Db	25903	GTAATGATGCCGGGGCAACAGTGAACCCAGGCACATGATGTCCCCACCCCAAGGCTTG	25962	
QY	108	-----	108	
Db	25963	GCCCTGACCTCTAGACATGTGGCCAGGGTGTGTGAAGATCTGCGCGGCCCTTTCACCC	26022	
QY	108	-----	108	
Db	26023	CAGTAAAGAGAGAGCTCCAGCGTATGGCCCCCACCACCAACAGAGATCCAGACCAAGCCA	26082	
QY	108	-----	108	
Db	26083	GGCAGCAGCAGGCGTTCAATGGCTGGCAGAGAGAGTACTCTGTGAGACGGGGGATTTAG	26142	
QY	108	-----	108	
Db	26143	GGATGGAGCGTTGGAGAGAGTGTAGGTGGAGACAGGCGCAAGGCTGTGAGAAATGGGG	26202	

QY 108 ----- 108
Db 26203 TCATTTAGAGTACCCCTTTCTTTCTATGTGGTCAAGTGTCTAGCCTCCAGGTGCAGGA 26262
QY 108 ----- 108
Db 26263 GCCATCCCTGGCGTCCAGGAAGACCTTGCTCCCTTTCTGAGGGCGCGCCCTCA 26322
QY 108 ----- 108
Db 26323 TTGACTGGTCCACCTTTCTGCTTGGCTGTGGCAGTTTCTCTGGCATCTGCTTTC 26382
QY 108 ----- 108
Db 26383 AGGAGCATGTCTCAGGCCCAFTTCAGATGAGAAGATGGGCTTCTGTTCCCGAGAGGGGT 26442
QY 108 ----- 108
Db 26443 GGTGCCAGCCTTTTCTGCCCTTCACGACCTCAGGCTCCATTTGCCAGTGATTTCTCAGCAG 26502
QY 108 ----- 108
Db 26503 ATCTCACACAGGGGGAAGGTGTCAGTGTACCAAGTGTCTCTGGGCTGTGGGGTTTGA 26562
QY 108 ----- 108
Db 26563 CAGAAGCTCCAGAAAGAAGTTGTAGTAATCATACAGCTGTCTATAGGCTCGCAGTTT 26622
QY 108 ----- 108
Db 26623 CTCTGAGCAGTTGCCTTGCTGAGTGCCAGGTTGGACTGTCTTACCTAAACCCCTGGAACAA 26682
QY 108 ----- 108
Db 26683 CACTTAACCTCCGTTGTGTGAGACTGAGGGGCGCAGAGGTCAACACAGCTGGGCGCTACGC 26742
QY 108 ----- 108
Db 26743 CAAGCCTTTGCTCTTAGCTTCGGCACTGTCTCTGCTGCCAGCGTCTCAGTGTGCTGCC 26802
QY 108 ----- 108
Db 26803 TCTCAGGGTGCACACCCCTTCGAGGCTGAAGGCTTTGTCAAAGACATTGATCTGCAGG 26862
QY 108 ----- 108
Db 26863 AAGCAGCGGTGCCTTGCAGTGCTGGGCCAGAGCAGGACTGACATGTGAAAGTCCAT 26922
QY 108 ----- 108
Db 26923 CACACAAGAACCTGCAGTGAAGACAGCAAGCTGCTGCTGATTAAATAGAGGACATTTT 26982
QY 109 ----- Ala 109
Db 26983 GGCAACCACTGACAGTGGTGTTTTGAACAGCCCTTGCCCTGTCTGTGTCCCTAGGCC 27042
QY 110 PheAspValValGluArgSerPheLeuGluSerIleAspAlaLeuAlaGluLysAla 129
Db 27043 TTCGATGTGGTGGAGAGGAGCTTCTCGAGTCCATTGACACGCTTGGCTGTGAAAGGCA 27102
QY 130 SerLeuGlnSerGlnLeuProGlu----- 137
Db 27103 AGCCTCCAGTCGCAATTTGCCAGA--GGTAAATTTCCCGAGCCGACACCCAGGGAGTCAAGT 27161
QY 137 ----- 137
Db 27162 CCAGGCCCAAGCTTTGCAAGGAGCATGACTCATCTACTTTCTTGACATTACTGGGCCAGA 27221
QY 137 ----- 137
Db 27222 GCAACAGCGTTAGGAGCAGTTCTCTGATGGTGACACTGGTGGTGGCCACAGGTGAGG 27281

QY 137 ----- 137
Db 27282 GACCTCGTCTCTCTGTGGCAGCTGCATGCTAAAGCTCAGCAGTGTCTGGGCAGTGG 27341
QY 137 ----- 137
Db 27342 TGGACATGAGGAAGGGCTTTTCCAGTCTTCTGCTCTGAATGGGTCCCCGCCCTTGTCAA 27401
QY 137 ----- 137
Db 27402 CCACCTGTCAATCGCCTACTGTCCACCTGCAGCGCTTACCTCTGTGTGTGAGTTCGGT 27461
QY 137 ----- 137
Db 27462 CATCTCCAGCTTTCTGGAAGAACCTTTAGCCTGGAGCTGTAGCGAAGGCTTCATCTTGCT 27521
QY 137 ----- 137
Db 27522 GTGCCATGACCATCTGTCTGTCTGTCCCGGCCCTCGGACTTTTGTCTCTTAAGACA 27581
QY 137 ----- 137
Db 27582 AAGGAGGCCGGGCATGTGTGCTCAGCCTGTAAATCCAGCAGCCTTGGGAGCGCGAGCG 27641
QY 137 ----- 137
Db 27642 GCGGATCACAGGTCAGGAGATCAAGACCATCCTGGGCTTAACATGGTGAACCCCGTCTC 27701
QY 137 ----- 137
Db 27702 TACTAAATAACAAAAATTTAGCTGGCGTGTGTGGCGCTTGTAGTCCAGCTACT 27761
QY 137 ----- 137
Db 27762 CTCGAGGCTGAAGCAGAGAGAATGTGTGAACCTGGGAGCGGAGCTTGTAGTGAGCCGAG 27821
QY 137 ----- 137
Db 27822 ATCATGCCACTGCACCTGTAGCCTGGGGACACAGCGAGACTCCATCTCAAAAAA 27881
QY 137 ----- 137
Db 27882 AAGAGTAAAGGAGTGGAGAGGCTTAGGAAGATGTCAGAGTTCTGTCTGCCCTGACC 27941
QY 138 ----- GlyValProGlnHisGlnLeuProGlnTyrGlnLysIle 151
Db 27942 CTCTGTTGATGGTTGTAGGGAGTCCCTCAGCACCAGCTGCCTCTCAGTATCAGAAGATC 28001
QY 152 LeuGluArgLeuLysThrLeuGluArgGluIleSerGlyGlyAlaMetAlaValValAla 171
Db 28002 CTTGAGAGACTCAAGACGTTAGAGAGGGAATTTTCGGAGGGGCCATGGCCGTTGTGGCG 28061
QY 172 ValLeuLeuAsnAsnLysLeuTyrValAlaAsnVal----- 183
Db 28062 GTCCCTTCTCAACAACAAGCTCTACGTGCCAAATGTGCGTGAGCCCCCTCTGTCCCAGGG 28121
QY 183 ----- 183
Db 28122 CAGGAGGACTGGGAGAGGTCAAGCCACAGGGGTGCGTGCATTTATTGACAATCTGCTTT 28181
QY 183 ----- 183
Db 28182 CCAGACACTTCAGCAGCTTTAAACCCAGGGTCTCTGTAGACCGTTGGGTATGTCCCTCTC 28241
QY 183 ----- 183
Db 28242 CACAGTGACGCTCAGTCCCAACTGGAAGGAGAAAGACGGGATGGGAGACAGGTGCC 28301
QY 183 ----- 183
Db 28302 TGGCCTTTAGTCCCTAATTTCTGCTTTTAACTCATTTCTGGACAGTCTGATCCCATCTTGG 28361
QY 183 ----- 183

Db 28362 GTGGCCTTGGGGCGCGCTCTTGTGAGAGGTTTACTGGCCATGATAATACCAACAGTA 28421
Qy 183 ----- 183
Db 28422 TCCCATGAAACAGCTCCGAGATGGGGCTAGTGACAGGACATGGGGTTTGTGAAGA 28481
Qy 183 ----- 183
Db 28482 CCAGACAGGTGACAGTTTCATAGTAAAGACCTGTGAGAGACCCTTGTGATGCTGCC 28541
Qy 183 ----- 183
Db 28542 TTTTAACTCTCAGTGAAGAAATTTCCACATATAGAACATAGAGTACTGATATT 28601
Qy 183 ----- 183
Db 28602 TAATGAACATTCATGGCCCATCCCAATCCGACAGTTATCACTGTGGCCAGCCTTCT 28661
Qy 183 ----- 183
Db 28662 TTGTCCCAAGTTCTCTATTATTGAACAAATGCCAGACATGATCATGTATCTGTAAA 28721
Qy 183 ----- 183
Db 28722 TAATTCAGTCTGATCTAGAGAGAGGTCTGTTTATTAAGATCATATCTTATGATT 28781
Qy 183 ----- 183
Db 28782 ACACGTAAAAAGTTAAACCAATTTCTTGGTATCACCAGCTGCTTAAGTGTACTCA 28841
Qy 183 ----- 183
Db 28842 CCCCCTGGTGTGAGCCTGTTTGTCCCATTTTCAGTATTTCCATGTGTGAATGCTGCC 28901
Qy 183 ----- 183
Db 28902 TTTTCCCTCTGCTGCTCCCGGTATGCCATATTTCTCTGTGTGATGCTTTGCTTAC 28961
Qy 184 -----
-GlyThrAsnArgAlaLeuLeuCysLysSerThrValAspGly 197
Db 28962 TGTTCACATTTCTGCCACAGGTACAAACCGTGACCTTTTATGCAATCGACAGTGGATGG 29021
Qy 198 LeuGlnValThrGlnLeuAsnValAspHisThrThrGluAsnGluAspGluLeuPheArg 217
Db 29022 TTGCAGGTGACACAGCTGAACGTGACACACACACAGAGAACGAGGATGAGCTTCCGT 29081
Qy 218 LeuSerGlnLeu----- 221
Db 29082 CTTTCCAGCTGGGTAGTGGGAGAGTGGGAGCGGAAGCTGATCCCATGGGCTCACCC 29141
Qy 221 ----- 221
Db 29142 TTGGCCTGCTTTGTGTGGGTAGAGAGCGTGTGTAGAGGGGCTGTGATCTTGGGC 29201
Qy 221 ----- 221
Db 29202 TCCCAAGCCAGCTGCTGGGGTTTCATCCACAGACTGCCGCTTACTGTTGGTGAATTGG 29261
Qy 221 ----- 221
Db 29262 AACTTAAGTTTCTGTGCTCAGATCCCGCGTGTAAATGAGGATAAATGAACTTACC 29321
Qy 221 ----- 221
Db 29322 TCGGAGTGTAGAGAGATTCAATTAGTGCATGTGAATGCTTCTTGGAATGCTGGCACA 29381
Qy 221 ----- 221
Db 29382 CAGAGACACTTACTGTATTGTGTGAGACTGAAGAGCCAAAGAAGTCCAGGAGCCC 29441
Qy 221 ----- 221

Db 29442 AGCTGCTGTGAGCTGCTCCTTCCCTTCCAGTGAAGTCCAGGCAAGATGTGGGGCACAG 29501
Qy 221 ----- 221
Db 29502 GGATGGCAGGGAGACGGCAGAGCAAGCTGCTCCGTGCCAGTGTGCTTGAGAGACGGGC 29561
Qy 221 ----- 221
Db 29562 TCTGAGTGAAGTGTGTGATGAGGGCGTGGGACTAGAGACACCAGGACTTTGGGTGAG 29621
Qy 221 ----- 221
Db 29622 CTGCTTCCAGAGTCCCTGTGTGTGTGCTTTCATTCTTATTCAGTGGCTTATTGCT 29681
Qy 221 ----- 221
Db 29682 TCTTCCATGACTGTGTCTGTGCCCTTCTTTTGTCTCTTGTGAAACAAGAAGCAG 29741
Qy 222 -----
-GlyLeuAspAlaGlyLysIleLysGlnValGlyIle 223
Db 29742 GATTGTGCACCTGTTTCCCTCCGTAGGCTTGATGCTGGAAGATCAAGCAGGTGGGGAT 29801
Qy 223 eileCysGlyGlnGluSerThrArgArgIleGlyAspTyrLysValLysTyrGlyTyrTrp 253
Db 29802 CATCTGTGGCAGAGAGACACCGCGGATCGGGATTAAGGTTAAATATGCTACAC 29861
Qy 253 AspIleAspLeuLeuSer-----AlaAlaLysSerLysProIleIleAlaGluProIle 271
Db 29862 GGACATTGACCTTCTCAGTGAAGTGCAGGCCAGCTGTCCCTGTGCTTCAAAAGAAGCA 29921
Qy 271 uileHisGlyAlaGlnProLeuAspGlyValThrGlyPheLeuValLeuMetSerGlu 291
Db 29922 A-----GGTCTTGAAGAGGCCAAGATGGAGATTAATGATGCTG-----GG 29963
Qy 291 Y-LeuTyrLysAlaLeuGlnAlaAlaHisGly-----ProGlyGlnAlaAsn 307
Db 29964 AGCTCAGAGCTGCAGTGAAGCCGTGATCATGCTGCACCTGCAGCTTGGGCAACAGCA 30023
Qy 307 IngIuIleAla-----AlaMetIleAspThrGluPheAlaLysGlnThrSerLeu 324
Db 30024 AGACCTGTCTCAAAAAAAGAGGTCCAGAAAGCTGGGCGACAGCG---AGACTGT 30080
Qy 324 spAlaValAlaGlnAlaValValAspArgValLysArgIleHisSerAspThrPheLeu 344
Db 30081 AGGCCCAAGTGCACATGTGCAGTCAAAAGTTCACAGGCCAGCAAACTCTTCTTCC 30140
Qy 344 ergLysGlyLysArg-----AlaArgPheCysProArgHisGluAspMetThrLeu 361
Db 30141 TC---CCAGAGAGTGTGTTTGGGGCAGGTTAAGCCCAAGCCAGCGAGGTGGTACTG 30197
Qy 361 euValArgAsnPheGlyTyrProLeuGlyLysMet-----SerGlnProThrPro--- 377
Db 30198 TAGAAACAAGGGGTGAGTGTAGCTGGCCAAAGCTAGAGGCAACATGATCCCAAG 30257
Qy 378 -----SerProAlaProAlaAlaGlyArgValTyrProValSer-----V 392
Db 30258 GGAAGCAGGCCAGCAAGCCAGTGGCATGGAGTGAAGACACAGCCCTGTGTGGTGTCT 30317
Qy 392 alProTyrSerSerAlaGlnSerThrSerLysThrSerValThrLeuSerLeuValMetP 412
Db 30318 GCCCATGGCTGCAGCAGAG-----TGTGAGATGAAATGTGTATTTC 30359
Qy 412 roSerGln-----GlyIleMetValAsn 420
Db 30360 CCTCTGAGAGACCTCTCTGACATCACCCAGCGCTCATGGCAGAGAAATGGGAGTGC 30419
Qy 420 LysAlaHisSerAlaSerThrLeuAspGluAlaThrProThrLeuThrAsnGlnSerProT 440
Db 30420 ATCATCACACGGCCAAAGCCCGTGCAGAGCTGTGAGTGAATCTTTTATGATTCCTTA 30479
Qy 440 hrlLeuThrLeuGlnSerThrAsnThrHis--ThrGlnSerSerSerSerSerAspGly 459
Db 30480 CTCAGGCCCTGTAGGCAAGTCACTGCATCACCCCAAGGCAATTTTCATTAAGAGAACT 30539

QY 460 GlyLeuPheArgSerArgProAlaHisSerLeuProGly 473
 Db 30540 GAGCATGGAGATCATCAGGACGCCAC-----CCTGGT 30572
 RESULT 13
 AC127784/C
 LOCUS
 DEFINITION Rattus norvegicus clone CH230-131B6, *** SEQUENCING IN PROGRESS
 ***, 59 unordered pieces.
 AC127784 176665 bp DNA linear HTG 19-JUL-2002
 AC127784.1 GI:21908163
 HTG: HTGS_PHASE1.
 Rattus norvegicus
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 176665)
 Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,
 Alsbrooks, S.L., Anaratunge, H.C., Are, J.R., Ayele, M., Banks, T.,
 Barbaria, J., Benton, J., Bimage, K., Blankenburg, K., Bonnin, D.,
 Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,
 Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
 Cartron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
 Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
 Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
 Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
 Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
 Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
 Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
 Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
 Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
 Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
 Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
 Hernandez, O., Hodgson, A., Hognes, M., Holloway, C., Hollins, B.,
 Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
 Jacobson, B., Jia, J., Johnson, R., Jolivet, S., Joudah, S.,
 Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
 Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
 Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louissegh, H.,
 Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
 Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
 Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,
 Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
 Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
 Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwuonu, G.,
 Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
 Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
 Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savery, G.,
 Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I.,
 Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,
 Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
 Tansey, J., Taylor, C., Taylor, T., Teifrod, B., Thomas, N., Thomas, S.,
 Umami, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
 Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K.,
 Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinstock, G., and gibbs, R.
 Direct Submission
 Unpublished
 2 (bases 1 to 176665)
 Worley, K.C.
 Direct Submission
 Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information

Center project name: GZWK
 Center clone name: CH230-131B6
 ----- Summary Statistics
 Sequencing vector: Plasmid;
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 120196 bases at least Q40
 Consensus quality: 127192 bases at least Q30
 Consensus quality: 131862 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 59 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1
 1032: contig of 1032 bp in length
 1132: gap of unknown length
 1133: contig of 1307 bp in length
 2439: gap of unknown length
 2539: gap of unknown length
 2440: contig of 1071 bp in length
 3610: gap of unknown length
 3611: contig of 1245 bp in length
 4955: gap of unknown length
 4956: contig of 1993 bp in length
 5056: gap of unknown length
 7048: gap of unknown length
 7148: gap of unknown length
 8902: contig of 1754 bp in length
 9002: gap of unknown length
 10567: contig of 1565 bp in length
 10667: gap of unknown length
 12258: contig of 1591 bp in length
 12358: gap of unknown length
 14216: contig of 1858 bp in length
 14316: gap of unknown length
 14317: contig of 1444 bp in length
 15760: gap of unknown length
 15860: gap of unknown length
 16935: contig of 1075 bp in length
 17036: gap of unknown length
 18984: contig of 1949 bp in length
 18985: gap of unknown length
 20295: contig of 1211 bp in length
 20395: gap of unknown length
 21672: contig of 1277 bp in length
 21772: gap of unknown length
 23061: contig of 1289 bp in length
 23161: gap of unknown length
 24596: contig of 1435 bp in length
 24696: gap of unknown length
 26381: contig of 1685 bp in length
 26481: gap of unknown length
 29059: contig of 2578 bp in length
 29159: gap of unknown length
 30322: contig of 1163 bp in length
 30422: gap of unknown length
 32594: contig of 2172 bp in length
 32694: gap of unknown length
 34702: contig of 2008 bp in length
 34802: gap of unknown length
 37016: contig of 2214 bp in length
 37116: gap of unknown length
 38321: contig of 1205 bp in length
 38421: gap of unknown length
 40849: contig of 2428 bp in length
 40949: gap of unknown length
 43133: contig of 2184 bp in length
 43233: gap of unknown length
 45386: contig of 2153 bp in length
 45486: gap of unknown length
 48376: contig of 2890 bp in length

* 48377 48476: gap of unknown length
* 48477 51407: contig of 2931 bp in length
* 51408 51507: gap of unknown length
* 51508 53492: contig of 1985 bp in length
* 53493 53592: gap of unknown length
* 53593 55777: contig of 2185 bp in length
* 55778 55877: gap of unknown length
* 55878 58654: contig of 2776 bp in length
* 58654 58753: gap of unknown length
* 58754 61259: contig of 2506 bp in length
* 61260 61359: gap of unknown length
* 61360 63193: contig of 1834 bp in length
* 63194 63294: gap of unknown length
* 63294 66481: contig of 3188 bp in length
* 66482 66581: gap of unknown length
* 66582 68046: contig of 1465 bp in length
* 68047 68146: gap of unknown length
* 68147 72431: contig of 4285 bp in length
* 72432 72531: gap of unknown length
* 72532 74844: contig of 2313 bp in length
* 74845 74944: gap of unknown length
* 74945 77951: contig of 3007 bp in length
* 77952 78051: gap of unknown length
* 78052 81664: contig of 3613 bp in length
* 81665 84688: gap of unknown length
* 84689 84788: contig of 2924 bp in length
* 84789 87987: gap of unknown length
* 87988 88087: contig of 3199 bp in length
* 88088 91325: gap of unknown length
* 91326 91425: contig of 3238 bp in length
* 91426 94907: gap of unknown length
* 94908 95007: contig of 3482 bp in length
* 95008 95007: gap of unknown length
* 95008 98189: contig of 3182 bp in length
* 98190 98289: gap of unknown length
* 98290 102290: contig of 4001 bp in length
* 102291 102390: gap of unknown length
* 102391 105219: contig of 2829 bp in length
* 105220 105319: gap of unknown length
* 105320 108533: contig of 3214 bp in length
* 108534 108634: gap of unknown length
* 108634 112138: contig of 3505 bp in length
* 112139 112239: gap of unknown length
* 112239 116139: contig of 3901 bp in length
* 116140 116239: gap of unknown length
* 116240 121227: contig of 4988 bp in length
* 121228 121327: gap of unknown length
* 121328 125446: contig of 4119 bp in length
* 125447 125546: gap of unknown length
* 125547 128869: contig of 3423 bp in length
* 128970 129069: gap of unknown length
* 129070 133782: contig of 4713 bp in length
* 133783 133882: gap of unknown length
* 133883 141065: contig of 7183 bp in length
* 141066 141165: gap of unknown length
* 141166 147194: contig of 6029 bp in length
* 147195 147294: gap of unknown length
* 147295 154395: contig of 7101 bp in length
* 154396 154495: gap of unknown length
* 154496 160801: contig of 6306 bp in length
* 160802 160901: gap of unknown length

Alignment Scores:

Pred. No.: 4.32e-17 Length: 176665
Score: 457.00 Matches: 158
Percent Similarity: 23.14% Conservative: 4
Best Local Similarity: 22.57% Mismatches: 11
Query Match: 17.71% Indels: 528
DB: 2 Gaps: 3

US-09-830-144-4 (1-504) X AC127784 (1-176665)

QY 12 GUGlNGlNPrOSeTTrPhrAspAspLeuProLeuCysHisLeu-SerGlyValGlySe 31

||||:|||||
Db 124370 GAGGAGCAGCCAGCTGGAGAGATGACCTGCTGCTCTCGCGGGTGGTTC 124311
QY 31 rAlAsErAsArSeTyTSeRAlAspGlyThylGlySerHisProProGluAs 51
Db 124310 AGCTTCGGGGGAGGACTACTCTGTATGSCAG-GGCAGGAGAGTCACTCTCGAGAGA 124252
QY 51 pSeTTrPLeu----- 54
Db 124251 CAACTGGCTAGTTCAAGTAAGTGGGTGGCTGAGCANAGAGTCATGAGCAAG 124192
QY 54 ----- 54
Db 124191 AAGTACTGTCTGTACTGATGATGAGCTTGCTGTGATGAGGCCATGGAA 124132
QY 54 ----- 54
Db 124131 GTGTGACGTGTGGGACAGGGGTGGACGAGCTGGAGGAAGATTCGTGTGAGATG 124072
QY 54 ----- 54
Db 124071 ATGGACCCAGTGCAGGGTCCAGATGTGTCTGAACCTTCTGTGTAGTCCCTCC 124012
QY 54 ----- 54
Db 124011 CCAATTCTTCTCTTCTTCTCTTACCACTGCATGCCCTGTTAGCCCTCCCTCC 123952
QY 55 -----LysPheArSeRgLyAsnAsnCySPhLeuTyrgLyAlPheAsnGlyTyra 72
Db 123951 ACCCTCCCTGTGGAGAGAGAAACAAGTCTCTCTGTATGGGGTCTTCAATAGGCTTAG 123892
QY 72 spGlyAsnArGValThrAsnPhValAlaGlnArgLeuSerAlaGlyLeuLeuGlyG 92
Db 123891 ATGGCAACCCGCTGACCAACTTGTGGCGCAGAGACTCTGCAGAGCTCTGCTGGGCC 123832
QY 92 lnuEuAsnAlaGlyHisAlaGlyAlaAspValArgArgValLeu----- 106
Db 123831 AGCTCAACACCGAACACACAGATGTGATGTGCAGGGGTCTGCTGCAGATAGAGAT 123772
QY 106 ----- 106
Db 123771 TGGGCTAGTGACAAACAGCTGACTGTGTGTCTTCACTCCCTGGAACTTGGGCC 123712
QY 106 ----- 106
Db 123711 CTGCATCTGTAAGTAGACAGAGCTGATGAAGTGTCTTCTTATCCCTTACCA 123652
QY 106 ----- 106
Db 123651 AGATCATTTTGGGGTACACTTCATGTGTGGCAGGGGCTCTGGCCACAGAGTCCAGGAC 123592
QY 106 ----- 106
Db 123591 AGGAGGAGAGAGAGAGCTTTTCATTGTAGACAGACTGACTTGGAGAAAGGAGAGACTG 123532
QY 106 ----- 106
Db 123531 GGAACAGAGCTGTGTAGATAGATACCTTTGTTGAGATTCTTATTTCTTCTTAC 123472
QY 106 ----- 106
Db 123471 CTAGGAAGTTTCTTGTGTTTGAATAATTGGCCCTAGATTCTTCTGTGTGTC 123412
QY 106 ----- 106
Db 123411 TCAGTGCTTAGACTGAGCTGTGCAGAGGCTAGACCAAGAGAAAGCTTCTGTCTTAC 123352
QY 106 ----- 106
Db 123351 CAGTGTCAAGCAGCTCTGTATCAAGGCCGAGAGTTGTATGCCCCCTACAGACTCAGTAC 123292
QY 106 ----- 106

Db 123291	CAAGCACTCTCGACTGTCTCACTTCACATGAACCCAAATGTGGAGCACTGCTGTTCA	123232
QY 106	-----	106
Db 123231	CCCAACACTCTTGAGCGCAGGGCTGGCCCTTCTCTCTCATTCCTTGAGGCTGCATAA	123172
QY 106	-----	106
Db 123171	GTAGCAACCTTTTCCGTCCTCCACAGGTAAGGTCTTTGTCAAACAGTGGACAGAAAGCA	123112
QY 106	-----	106
Db 123111	CACATATAGGCAGAAATGATATGTCACTGTCAACAGAGCCATCCATGAATGCATGG	123052
QY 106	-----	106
Db 123051	ATCTGTTACTCTCCATAGGACATCTTAACAGAGGTGATGGCTTATTTGAGCCAACCC	122992
QY 107	-----LeuGlnAlaPheAspValValGluArgSerPheLeuGluSer	120
Db 122991	TTGCCTTGCCCTGTGTCTCTTCAGGCTTCGATGTGGTGAGAGGAGCTTCTCTGGAGTCT	122932
QY 121	IIeAspAlaLeuAlaGluLysAlaSerLeuGlnSerGlnLeuProGlu	137
Db 122931	ATCGATGATGCCCTTGGCTGAGAAAGCAAGCTCCAGTCCAGCTGCCGAGGTAACCGAC	122872
QY 137	-----	137
Db 122871	TGTCGTACTAGTACTGCTGGGTGCCAGCACTCACCTCTGTATCATCTTGCACCCACAGG	122812
QY 137	-----	137
Db 122811	CACAGTAAGGCTGCAGCAGGAATCATGAGCAGCATCTCTGATGGCAGGCTGGGGTGGC	122752
QY 137	-----	137
Db 122751	ACTGGGTATGTCCCTGTAGAGACTCCAGCCACAGTGGGCGACGTTTTCTCGAGGCTCA	122692
QY 137	-----	137
Db 122691	TCCTTAGATGGCAGGATAGGTGAAGGAAGGCTGTCCAGGTTCTGTGATCTGTTCTAG	122632
QY 137	-----	137
Db 122631	TCCAGAGCCAGGCTCACCTTCAGGCTGCCCTAGTTAAAGCAACAGTACCTTCCACGG	122572
QY 137	-----	137
Db 122571	CAGCCTACAGGTATCCAGCCTGGCCACACCTATGTTGTCATCTGTCTTGGCTGCTCAA	122512
QY 137	-----	137
Db 122511	GGGCCTTGATATCTTTATCTAAAGGAAAGAGAGGATGTTAGGTGGTACTCAAGCCCCC	122452
QY 138	-----GlyValProGlnHisGlnLeuProPro	146
Db 122451	ATGATGCTCTGATCTCTTGATGATGTTTTCAGGGGTGTCCTCCCAACACAGCTGTACT	122392
QY 147	GlnTyrGlnLysIleLeuGluArgLeuLysThrLeuGluArgGluIleSerGlyGlyAla	166
Db 122391	CAGTATCAGAAGATCTCTGGAGAGACTCAAGGCACCTGGAGAAGGAGATCTTCGGGAGG	122332
QY 167	MetAlaValValAlaValLeuLeuAsnLysLeuTyrValAlaAsnValGly	184
Db 122331	ATGGCTGTTGGCGGTCTCTTCTCAACAACAAGCTCTACGTTGCCAATGTTGGT	122278
RESULT 14		
AC127784	176665 bp DNA linear HTG 19-JUL-2002	
LOCUS	Rattus norvegicus clone CH230-131B6, *** SEQUENCING IN PROGRESS	
DEFINITION	***, 59 unordered pieces.	
ACCESSION	AC127784	
VERSION	AC127784.1 GI:21908163	

HTG; HTGS_PHASE1.
Rattus norvegicus.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 176665)
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C., Alsbrooks, S.L., Amaraturge, H.C., Are, J.R., Ayele, M., Banks, T., Barbara, J., Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureahi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseged, H., Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, B., Nwokwenkwo, S., Oguh, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I., Sodergren, E., Sonaik, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Umami, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Woodson, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstock, G. and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 176665)
Worley, K.C.
Direct Submission
Submitted (19-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information

Center project name: GZWK
Center clone name: CH230-131B6
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 120196 bases at least Q40
Consensus quality: 127192 bases at least Q30
Consensus quality: 131862 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

NOTE: This is a 'working draft' sequence. It currently consists of 59 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1	1032: contig of 1032 bp in length
1033	1132: gap of unknown length
1133	2439: contig of 1307 bp in length
2440	2539: gap of unknown length
2540	3610: contig of 1071 bp in length
3611	3711: gap of unknown length
3711	4955: contig of 1245 bp in length
4956	5055: gap of unknown length
5056	7048: contig of 1993 bp in length
7049	7148: gap of unknown length
7149	8002: contig of 1754 bp in length
8903	9020: gap of unknown length
9003	10567: contig of 1565 bp in length
10568	10667: gap of unknown length
10668	12258: contig of 1591 bp in length
12259	12258: gap of unknown length
12359	14216: contig of 1858 bp in length
14217	14316: gap of unknown length
14317	15760: contig of 1444 bp in length
15761	15860: gap of unknown length
15861	16935: contig of 1075 bp in length
16936	17035: gap of unknown length
17036	18984: contig of 1949 bp in length
18985	19084: gap of unknown length
19085	20295: contig of 1211 bp in length
20296	20395: gap of unknown length
20396	21672: contig of 1277 bp in length
21673	21772: gap of unknown length
21773	23061: contig of 1289 bp in length
23062	23161: gap of unknown length
23162	24596: contig of 1435 bp in length
24597	24696: gap of unknown length
24697	26381: contig of 1685 bp in length
26382	26881: gap of unknown length
26882	29059: contig of 2578 bp in length
29060	29159: gap of unknown length
29160	30322: contig of 1163 bp in length
30323	30422: gap of unknown length
30423	32594: contig of 2172 bp in length
32595	32694: gap of unknown length
32695	34702: contig of 2008 bp in length
34703	34802: gap of unknown length
34803	37016: contig of 2214 bp in length
37017	37116: gap of unknown length
37117	38921: contig of 1205 bp in length
38922	38821: gap of unknown length
38822	40849: contig of 2428 bp in length
40850	40949: gap of unknown length
40950	43333: contig of 2184 bp in length
43134	43333: gap of unknown length
43234	45886: contig of 2153 bp in length
45887	45886: gap of unknown length
45887	48376: contig of 2890 bp in length
48377	48476: gap of unknown length
48477	51007: contig of 2931 bp in length
51408	51507: gap of unknown length
51508	53992: contig of 1985 bp in length
53493	53592: gap of unknown length
53593	55777: contig of 2185 bp in length
55778	55877: gap of unknown length
55878	58653: contig of 2776 bp in length
58654	58753: gap of unknown length
58754	61593: contig of 2506 bp in length
61260	61359: gap of unknown length
61360	63193: contig of 1834 bp in length

*	63194	63293:	gap of unknown length
*	63394	66481:	contig of 3188 bp in length
*	66582	66581:	gap of unknown length
*	66582	68046:	contig of 1465 bp in length
*	68047	68146:	gap of unknown length
*	68147	72431:	contig of 4285 bp in length
*	72432	72532:	gap of unknown length
*	72532	74544:	contig of 2133 bp in length
*	74845	74944:	gap of unknown length
*	74945	77951:	contig of 3007 bp in length
*	77952	78051:	gap of unknown length
*	78052	81664:	contig of 3613 bp in length
*	81665	81764:	gap of unknown length
*	81765	84688:	contig of 2924 bp in length
*	84689	84788:	gap of unknown length
*	84789	88087:	contig of 3199 bp in length
*	87888	88087:	gap of unknown length
*	88088	91325:	contig of 3238 bp in length
*	91326	91425:	gap of unknown length
*	91426	94907:	contig of 3482 bp in length
*	94908	95007:	gap of unknown length
*	95008	98189:	contig of 3182 bp in length
*	98190	98289:	gap of unknown length
*	98290	102290:	contig of 4001 bp in length
*	102291	102390:	gap of unknown length
*	102920	105219:	contig of 2829 bp in length
*	105220	105319:	gap of unknown length
*	105320	108633:	contig of 3214 bp in length
*	108634	108633:	gap of unknown length
*	108634	112138:	contig of 3505 bp in length
*	112239	112238:	gap of unknown length
*	112239	116139:	contig of 3901 bp in length
*	116240	116239:	gap of unknown length
*	116240	121227:	contig of 4988 bp in length
*	121228	121227:	gap of unknown length
*	121328	125446:	contig of 4119 bp in length
*	125447	125546:	gap of unknown length
*	125547	128969:	contig of 3423 bp in length
*	128970	129069:	gap of unknown length
*	129070	133782:	contig of 4713 bp in length
*	133783	133882:	gap of unknown length
*	133883	141065:	contig of 7183 bp in length
*	141066	141165:	gap of unknown length
*	141166	147994:	contig of 6029 bp in length
*	147195	147294:	gap of unknown length
*	147295	154495:	contig of 7101 bp in length
*	154396	154495:	gap of unknown length
*	154496	160801:	contig of 6306 bp in length
*	160802	160901:	gap of unknown length

Alignment Scores:	
Pred. No.:	1,98e-14
Score:	414.00
Percent Similarity:	25.17%
Best Local Similarity:	22.90%
Query Match:	16.05%
DB:	2
	4
Length:	17665
Matches:	131
Conservative:	13
Mismatches:	20
Indels:	408
Gaps:	4

[illegible]

```
|||||
Db 161880 AGGGGCGGTGTACCCCTGTTCTGTGCGCCCTACTCAAGTGCACGAGCACGACGAC 161939
QY 403 rSerValThrLeuSerLeuValMetProSerGlnGlyGlnMetValAsnGlyAlaHisSe 423
Db 161940 CAGTGTGACTCTGTCCTCTCATGCTTCTCAGGGCCAGATGGTCAAAGGCTCTCACAG 161999
QY 423 rAlaSerThrLeuAspGluAlaThrProThrLeuThr----- 435
Db 162000 TGCCCTCCACCTTGGATGAAGCCACTCTACTCTCACTAAGTAAGCTATCCTACTCTAGGT 162059
QY 435 ----- 435
Db 162060 ACCTGGAGAGGGACAAACCGCTGTACAGCCAAGCGCTAAGGCTCAGTAGGGTGGGTAAGA 162119
QY 435 ----- 435
Db 162120 GGCTGGGTGTGGGACCGTGGGTCTTTGACCTCCAATTCCCCCTTCCCTTTGGGTA 162179
QY 435 ----- 435
Db 162180 GGGAGTCTTATTAGACACAGGTAGGAAGGACAACTACTGTTAATTCTAGTAGTGTGA 162239
QY 435 ----- 435
Db 162240 CAGGACCTGGGGCAAGTGGAGGTTTCAGGGCAAGGTGGACAAGGGGCCGAGCAAGGAGC 162299
QY 435 ----- 435
Db 162300 AAGTGGCTGACTGATGTCCTGGGGCAGCTTTTCATCCATGGCAGCTTCTCTGGTTTGTC 162359
QY 435 ----- 435
Db 162360 CTTCTGTTTTGGAGCTGTAAACAGGGTCACTGGCTTCACGAATCTGTAGTTGGCCAGATT 162419
QY 435 ----- 435
Db 162420 CATTCCTGGCCTGTACTGGGTCTTCAGAGCTGGGCTCCCTCACTGTGACGGACCATTTG 162479
QY 435 ----- 435
Db 162480 CTGTCACTGCTAGTTTAAAGGAGCAGGAGGTAGAGGCTAGGGCATGGTCTGAGAGGGGCA 162539
QY 435 ----- 435
Db 162540 TGGTGCAGACCTGCTGCACAGGCTATTATCTACCGCCCCCTTCATTTCTGTGACTGTGT 162599
QY 435 ----- 435
Db 162600 ACTCTGCCCCCACCACCTCTCTCAGGGGGATCTGGTTTCCATTCACAGCCCTGCACGTG 162659
QY 435 ----- 435
Db 162660 GTGGAATAATACAGCTGCCCCACCAGCCCTCTTCAGGATAGAGCTTAGCCGATGTGG 162719
QY 435 ----- 435
Db 162720 CCAATTAGAGGGCTGGCAGGAGGAGTGTCCCTGTTTCTCTGCTCTTCCTTCACAGGGC 162779
QY 435 ----- 435
Db 162780 CTCCCATCTGTCCCAAGTCTTTGCTGCTGCTTTTCCACCACCACCTCTCTGGATAGA 162839
QY 435 ----- 435
Db 162840 GGGCCTGGTGTAGACATCTGGGCATCATGGGTAGCGGTGTGGTGTCCCCCAGGACAGTCAA 162899
QY 435 ----- 435
Db 162900 CATAGCTCCCCCAACCCCCACCATGCTTTCCCCAGGGAACGCCAGCTGTGTCCACACA 162959
QY 435 ----- 435
```

```
Db 162960 GCCAACCCAGGCAAACTTCATGCTTCTGTGTGACTGCTGCTGAACCTCACAGGTCTCTTCC 163019
QY 435 ----- 435
Db 163020 TCTGGCTCTGAGCCTTATTCTGTACCTAGTCAAGCCTGCTTTTAAACCATCTCACCC 163079
QY 435 ----- 435
Db 163080 CACTGGCTTTTATGATATTTGGCAGAGAGGCCTGAGCCTAAAGGGCCACTCTTTTCTCACT 163139
QY 435 ----- 435
Db 163140 GGGTGGTGGGAAGTTGACGGCCCTTCAGGCAGCTACCCAGACTGATTGCTCCCTCTCT 163199
QY 436 -----AsnGlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerS 454
Db 163200 TCTGTAGCCAGAGCCCACTCTGACCTGACCTGACCTACCAACCCACACCCAGAGCAGCA 163259
QY 454 erSerSerSerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProProGlyG 474
Db 163260 GCTCCAGCTCGACGGGGCCCTCTTCGCTCCAGACCGGCTCACTCACTTCCACCTGGAG 163319
QY 474 luAspGlyArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerVala 494
Db 163320 AGGATGGCAGGCTCGAGCCCTATGTGACTTTGTGAGTTTACCGACTCTGGAGCGTGG 163379
QY 494 spHisGlyGluGlnSerValValThrAlaPro 504
Db 163380 ACCATGGCAGCAGAGTGTGATACGGCACCT 163411
RESULT 15
LOCUS AC127924 141498 bp DNA linear HTG 19-JUL-2002
DEFINITION Rattus norvegicus clone CH230-23318, *** SEQUENCING IN PROGRESS
AC127924
AC127924.1 GI:21908451
VERSION HTG: HTGS PHASE1.
KEYWORDS Rattus norvegicus.
SOURCE Rattus norvegicus.
ORGANISM Rattus norvegicus.
REFERENCE 1 (bases 1 to 141498)
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Devila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jollivet,S., Joudan,S.,
Karleson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,C., Liu,J., Liu,W., Louleghed,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Mosser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

Peters, L., Pickens, R., Primus, E., Pu, L., Qules, M., Ren, Y.,
Rivas, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G.,
Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I.,
Sodergren, E., Sonaite, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, J., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Usmani, K., Vaequez, L., Vera, V., Villalon, D., Vinson, R., Wang, O.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S.,
Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 141498)
Worley, K.C.
Direct Submission
Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
----- Genome Center of Medicine
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KXAS
Center clone name: CH230-23318
----- Summary Statistics
Sequencing vector: Plasmid:
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 86446 bases at least Q40
Consensus quality: 92843 bases at least Q30
Consensus quality: 96514 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 68 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1030: contig of 1030 bp in length
1130: gap of unknown length
1131: contig of 1132 bp in length
2263: gap of unknown length
2362: gap of unknown length
3440: contig of 1078 bp in length
3441: gap of unknown length
3540: gap of unknown length
5167: contig of 1627 bp in length
5168: gap of unknown length
5267: gap of unknown length
5268: contig of 1113 bp in length
6580: gap of unknown length
6581: gap of unknown length
6680: gap of unknown length
6881: contig of 1203 bp in length
7883: gap of unknown length
7884: gap of unknown length
7984: contig of 1038 bp in length
9021: contig of 1038 bp in length
9022: gap of unknown length
9122: contig of 1704 bp in length
10825: contig of 1704 bp in length
10826: gap of unknown length
10925: gap of unknown length
10926: contig of 1158 bp in length
12083: gap of unknown length
12183: gap of unknown length
12184: contig of 1127 bp in length
13110: gap of unknown length
13111: gap of unknown length
13410: gap of unknown length
13411: gap of unknown length
14464: contig of 1054 bp in length
14465: gap of unknown length
14565: gap of unknown length
16304: contig of 1740 bp in length
16305: gap of unknown length
16404: gap of unknown length
16405: contig of 1935 bp in length
18439: gap of unknown length
18440: contig of 1054 bp in length
19493: gap of unknown length
19494: gap of unknown length
19593: gap of unknown length

19594: contig of 1705 bp in length
21298: gap of unknown length
21399: contig of 1119 bp in length
21399: gap of unknown length
22537: gap of unknown length
22538: contig of 1492 bp in length
24129: gap of unknown length
24130: gap of unknown length
24230: contig of 1048 bp in length
25278: gap of unknown length
25278: gap of unknown length
25378: contig of 1701 bp in length
27079: gap of unknown length
27179: gap of unknown length
29508: contig of 2330 bp in length
29609: gap of unknown length
31421: contig of 1812 bp in length
31421: gap of unknown length
31521: contig of 1714 bp in length
33234: gap of unknown length
33235: gap of unknown length
33335: contig of 1776 bp in length
35110: gap of unknown length
35111: gap of unknown length
35210: gap of unknown length
36675: contig of 1465 bp in length
36675: gap of unknown length
36775: gap of unknown length
36776: gap of unknown length
37894: contig of 1119 bp in length
37895: gap of unknown length
37994: gap of unknown length
39970: contig of 1976 bp in length
40070: gap of unknown length
40071: gap of unknown length
41623: contig of 1553 bp in length
41624: gap of unknown length
41724: gap of unknown length
43157: contig of 1434 bp in length
43257: gap of unknown length
45044: contig of 1787 bp in length
45144: gap of unknown length
45145: gap of unknown length
46604: contig of 1460 bp in length
46704: gap of unknown length
46705: gap of unknown length
46705: contig of 1194 bp in length
47898: gap of unknown length
47999: gap of unknown length
49208: contig of 1210 bp in length
49309: gap of unknown length
49309: gap of unknown length
50615: contig of 1307 bp in length
50715: gap of unknown length
50716: gap of unknown length
52499: contig of 1784 bp in length
52599: gap of unknown length
52500: gap of unknown length
54003: contig of 1404 bp in length
54004: gap of unknown length
54103: gap of unknown length
54104: gap of unknown length
55661: contig of 1558 bp in length
55662: gap of unknown length
55761: gap of unknown length
58110: contig of 2349 bp in length
58110: gap of unknown length
58211: gap of unknown length
61322: contig of 3112 bp in length
61323: gap of unknown length
61423: gap of unknown length
61423: gap of unknown length
62808: gap of unknown length
62908: gap of unknown length
62909: gap of unknown length
65459: contig of 2551 bp in length
65460: gap of unknown length
65559: gap of unknown length
66882: contig of 1323 bp in length
66883: gap of unknown length
66982: gap of unknown length
68456: contig of 1474 bp in length
68457: gap of unknown length
68556: gap of unknown length
68557: gap of unknown length
70789: contig of 2233 bp in length
70790: gap of unknown length
70889: gap of unknown length
73103: contig of 2214 bp in length
73104: gap of unknown length
73203: gap of unknown length
73204: gap of unknown length
74721: contig of 1518 bp in length
74722: gap of unknown length
74821: gap of unknown length
77416: contig of 2595 bp in length
77417: gap of unknown length
77516: gap of unknown length
77517: gap of unknown length
78949: contig of 1433 bp in length
78950: gap of unknown length
79049: gap of unknown length
81501: contig of 2452 bp in length
81502: gap of unknown length
81601: gap of unknown length
81602: gap of unknown length
83009: contig of 1408 bp in length
83109: gap of unknown length
83110: gap of unknown length
84773: contig of 1664 bp in length
84874: gap of unknown length
84874: gap of unknown length
86891: contig of 2018 bp in length

* 86892 86991: gap of unknown length
* 86992 88782: contig of 1791 bp in length
* 88783 88882: gap of unknown length
* 88883 90589: contig of 1707 bp in length
* 90590 90689: gap of unknown length
* 90690 92680: contig of 1991 bp in length
* 92681 92780: gap of unknown length
* 92781 94508: contig of 1728 bp in length
* 94509 94608: gap of unknown length
* 94609 97262: contig of 2654 bp in length
* 97263 97362: gap of unknown length
* 97363 100119: contig of 2757 bp in length
* 100120 100219: gap of unknown length

Alignment Scores:

Pred. No.: 7,03e-13 Length: 141498
Score: 387.00 Matches: 121
Percent Similarity: 45.95% Conservative: 38
Best Local Similarity: 34.97% Mismatches: 97
Query Match: 15.00% Indels: 90
DB: 2 Gaps: 15

US-09-830-144-4 (1-504) x AC127924 (1-141498)

QY 211 AsnGluAspGluLeuPheArgLeuSerGlnLeuGlyLeuAspAlaGlyLysIleLysGln 230
Db 83252 TCTAGGACAGCGCTGTAGACTTGGAGGCTGTGGAGCTGGAGTTCGGAAGAGCCAGAT 83311
QY 231 ValGlyIleIleCysGlyGln-----GluSerThrArgArgIleGlyAsp 245
Db 83312 ACAAGATTATCTGCCGCCAGCGGTGTGCATGTGTGCACCTTACAGCTCAAGCCGCTCT 83371
QY 246 TyrLysValLysTyrGlyTyrThrAspIleAspLeuSerAlaAlaLysSer---Lys 264
Db 83372 TTCTTTCTAGA-----ACTGATTGGCTCAGGGAACAATTGCTAGGCTGTAGT 83422
QY 265 ProIleIleAlaGluPro-----GluIleHisGlyAlaGlnProLeuAspGlyValThr 282
Db 83423 CAGGCATTGGCTCTGCCTGTCTAGGCAGCTGCATTCCCTCC-----LysAlaLeuGluAla 298
QY 283 GlyPheLeuValLeuMetSerGluGlyLeuTyr-----LysAlaLeuGluAla 298
Db 83465 ---TGGGTCTATCTTACCTCGCCTCTCGCTCTATCTTCCAGGATCTCGATCTCTAAGTCTC 83521
QY 299 AlaHis-----GlyPro 302
Db 83522 CTACACAGGCTTTGGATCCCCCTTACCCGCCCTTCTCCTCATCAGGTTCT 83581
QY 303 GlyGlnAlaAsnGlnGluIleAlaAlaMetIleAspThrGluPheAlaLysGlnThrSer 322
Db 83582 CCCTCCACACCCAGGAGGTTGCGCAATGATTGACACCGAATTTGCCAAGCAGACCTCC 83641
QY 323 LeuAspAlaValAlaGlnAlaValAlaAspArgValLysArgIleHisSerAspThrPhe 342
Db 83642 CTCGATGCAGTTGCCAGGCTGTGTAGACCGGTAAAGCGTATCCACAGTGCACACTTT 83701
QY 343 AlaSerGlyGlyGluArgAlaArgPheCysProArgHisGluAspMetThrLeuLeuVal 362
Db 83702 GCCAGCGGTGGGAGCGGGCCAAAGTTCTGCCACGGCATGAAGACATGACCTTGTGGTG 83761
QY 363 ArgAsnPheGlyTyrProLeuGlyGluMetSerGlnProThrProSerProAlaProAla 382
Db 83762 AGGAACCTTTGGCTATCATTTAGGTGAATGAGCCAGCCGACACCCACCGGCCCGCCAGGT 83821
QY 383 -----AlaGlyGlyArg 386
Db 83822 ACGTATGCAATGTACTGTACAGGTAGAAGCACCTGAGATGAGTCTGGGCTGGGGCAGC 83881
QY 387 Val-----TyrProValSerVal 392
Db 83882 ACAACATGAGCTGACAAACATGAGCTCGCTGTTGGAATGCAGAGGAGCCAGTGGCCTGG 83941
QY 393 ProTyrSerSerAlaGlnSerThrSerLysThrSerValThrLeuSerLeuValMetPro 412

Db 83942 ---CTTTGCTCAACACAGCGTGAAGTCTGAGCTGCAAAATCTTTATTCGTTTCTG--- 83992
QY 413 SerGlnGlyGlnMetVal-----AsnGlyAlaHis 422
Db 83993 AAAGAAGGATCAATTGCTCTCTGGCCACCCCTGTTATCTTGTGTCATGAGCAGTGCAC 84052
QY 423 -----SerAlaSerThrLeuAspGluAlaThrPro---ThrLeuThr 435
Db 84053 AAAGCATGGTAGAGTGACACAGCTGCCACCATCGATGGTGGAGTCCCTGTTAGCCTATGG 84112
QY 436 AsnGlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 455
Db 84113 TCAGTGTGGTCACTTCTCACCTTTTACTAGCAACCCCTAACAGGAAGGAGAAAGGCTCTAGC 84172
QY 456 SerSerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGlyGluAsp 475
Db 84173 ATGAAGGACCTGGATGAGTCAAGTGTGGACAGATGGTCCACCCCA-----GAACCT 84226
QY 476 GlyArgValGluProTyr 481
Db 84227 GGCACCTGTCTGACCATTT 84244

Search completed: December 10, 2002, 01:12:13
Job time : 3382.68 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 9, 2002, 22:57:35 ; Search time 249.44 Seconds
(without alignments)
4550.213 Million cell updates/sec

Title: US-09-830-144-4

Perfect score: 2580

Sequence: 1 MAAQRRLQLSQSPQSWTDD.....AEFYRLWSVDHGEQSVVTAP 504

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-Q=/cgn2_1/USPTO.spool/US09830144/runat_04122002_141353_2242/app_query.fasta_1.1422
-DB=N_Geneseq_101002 -qfmt=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09830144 @CNC 1.1.113 @runat_04122002_141353_2242 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MWAP -LARGEQUERY -NEG_SCORES=0 -WAIT -JONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2580	100.0	1515	21	AAZ48861 Human TAB1 coding
2	2580	100.0	1560	18	AAZ91175 Human TAB1 (TAK1) b
3	2580	100.0	1560	20	AAZ56278 Human TAB1 encodin
4	2580	100.0	1560	21	AAZ39106 Human TAB-1 nucleo
5	2580	100.0	1569	20	AAZ56282 Human TAB1-FLAG en
6	2577	99.9	1568	20	AAZ56310 Human TAB1 encodin
7	2575	99.8	1560	18	AAZ91178 Human TAB1 (TAK1) b
8	543.5	21.1	16877	22	ABA20494 Human nervous syst
9	543.5	21.1	16877	22	AAL36984 Human musculoskele
10	409	15.9	696	21	AAF15895 Human prostate can
11	365	14.1	211	22	ABA47246 Human breast cell
12	365	14.1	211	22	ABA65131 Human foetal liver
13	365	14.1	211	22	ABA32233 Probe #10699 for g
14	365	14.1	211	22	AAK13550 Human brain expres
15	365	14.1	211	22	AAK39289 Human bone marrow
16	365	14.1	211	22	AAK120102 Probe #10035 for g
17	365	14.1	211	22	AAI45302 Probe #13988 used
18	365	14.1	211	22	AAI05808 Probe #5799 used t
19	365	14.1	211	24	ABG13376 Human genome-deriv
20	323.5	12.5	455	22	ABA42092 Human breast cell
21	323.5	12.5	455	22	ABA52514 Human foetal liver
22	323.5	12.5	455	22	ABA22303 Probe #769 for gen
23	323.5	12.5	455	22	AAK00778 Human brain expres
24	323.5	12.5	455	22	AAK26230 Human bone marrow
25	323.5	12.5	455	22	AAI10863 Probe #796 for gen
26	323.5	12.5	455	22	AAI32123 Probe #809 used to
27	323.5	12.5	455	22	AAI00787 Human genome-deriv
28	323.5	12.5	455	22	ABG00818 Human breast cell
29	271.5	10.5	467	22	ABA42093 Human foetal liver
30	271.5	10.5	467	22	ABA52515 Probe #770 for gen
31	271.5	10.5	467	22	ABA22304 Human brain expres
32	271.5	10.5	467	22	AAK00779 Human bone marrow
33	271.5	10.5	467	22	AAK26231 Probe #797 for gen
34	271.5	10.5	467	22	AAI10864 Probe #810 used to
35	271.5	10.5	467	22	AAI32124 Probe #779 used to
36	271.5	10.5	467	22	AAI00788 Human genome-deriv
37	271.5	10.5	467	24	ABG00819 Human breast cell
38	261	10.1	194	22	ABA47247 Human foetal liver
39	261	10.1	194	22	ABA65132 Probe #10700 for g
40	261	10.1	194	22	ABA32234 Human brain expres
41	261	10.1	194	22	AAK13551 Human bone marrow
42	261	10.1	194	22	AAK39290 Probe #10036 for g
43	261	10.1	194	22	AAI20103 Probe #13989 used
44	261	10.1	194	22	AAI45303 Probe #5800 used t
45	261	10.1	194	22	AAI05809

ALIGNMENTS

RESULT 1

AAZ48861

ID AAZ48861 standard; cDNA; 1515 BP.

XX AAZ48861;

AC AAZ48861;

XX 24-MAR-2000 (first entry)

DT Human TAB1 coding sequence.

DE Human; TAB1; XIAP; X-linked inhibitor of apoptosis protein; TGF-beta;

XX transforming growth factor-beta activated kinase 1; monocytic migration;

KW TAK1 binding protein 1; extracellular matrix protein production;

KW cell growth inhibitor; beta-amyloid protein deposition;

KW immunosuppression; Transforming growth factor-beta; ds.

XX Homo sapiens.

OS

XX

PN JP1326328-A.
 XX 26-NOV-1999.
 XX 13-MAY-1998; 98JP-0130378.
 XX 13-MAY-1998; 98JP-0130378.
 PA (MATS/) MATSUMOTO K.
 XX MPI: 2000-078337/07.
 DR P-PSDB; AAY59450.
 XX Screening a substance which inhibits combination of the X-linked
 PT inhibitor of apoptosis protein -
 PS Claim 2; Page 25-26; 43pp; Japanese.
 XX
 CC This sequence encodes the human TAB1 protein.
 CC The invention relates to a method for screening a substance inhibiting
 CC the formation of a complex between XIAP and TAB1, in which X-linked
 CC inhibitor of apoptosis protein (XIAP), transforming growth factor-beta
 CC activated kinase 1 (TAK1) binding protein 1 (TAB1) and a substance to be
 CC tested are contacted with each other and then the presence or formation
 CC of a complex between XIAP and TAB1 is detected. The substance can be used
 CC as a drug for extracellular matrix protein production enhancement, cell
 CC growth inhibition, monocyte migration, physiologically active substance
 CC induction, immunosuppression, and beta-amyloid protein deposition. A
 CC substance inhibiting the formation of a complex between TAB1 and XIAP as
 CC well as between XIAP and TGF-beta (Transforming growth factor-beta) type
 CC I and/or type II receptor is useful as a drug.
 XX
 SQ Sequence 1515 BP; 323 A; 457 C; 463 G; 272 T; 0 other;

Alignment Scores:

Pred. No.: 8, 65e-185 Length: 1515
 Score: 2580.00 Matches: 504
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 21 Gaps: 0

US-09-830-144-4 (1-504) X AAZ48661 (1-1515)

QY 1 MetAlaIaGlnIaRgaRgSerIeuleuGlnSerGluGlnProSerTrpThrAsp 20
 Db 1 ATGGCGGCGAGAGAGAGCTTGCTGCAGAGTGAACAGCCAGCTGGACAGATGAC 60
 QY 21 LeuProleuCyHisIeuserGlyValGlySerAlaSerAsnArgSerTrpSerIaAsp 40
 Db 61 CTGGCTCTTGCCACCTCTCTGGGGTTGGCTCAGCTCCAAACCGCAGCTACTGCTGAT 120
 QY 41 GlyLysGlyThrGluSerHisProProGluAspSerTrpLeuIysPheArgSerGluAsn 60
 Db 121 GGCAGAGGCACTGAGAGCCACCGCCAGAGAGACAGCTGGCTCAAGTTCAAGAGTGAAC 180
 QY 61 AsnCySpheLeuTrpGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnPheVal 80
 Db 181 AACTGCTTCTCTATGGGGCTTCAACGGCTATGATGGCAACCGATGACCAACTTCGTG 240
 QY 81 AlaGluArgIeuserAlaGluLeuLeuGluGlnLeuAsnAlaGluHisAlaGluAla 100
 Db 241 GCCCAGGCGCTGTCGAGAGCTCTGCTGGGCGCACTGAAATGCCAGACCGCCGAGGCC 300
 QY 101 AspValArgArgValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGluSer 120
 Db 301 GAGCTCGGGGTGCTGCTGCGAGGCTTGAGTGTGTGAGAGGAGCTTCTCGAGTCC 360
 QY 121 IleAspAspAlaLeuAlaGluLysAlaSerLeuGlnSerGlnLeuProGluGlyValPro 140
 Db 361 ATTGACGACCGCTTGCTGAGAGGCAAGGCTCCAGTCCCAATTGCGCAAGAGGAGTCCCT 420
 QY 141 GlnHisGlnLeuProProGlnTrpGlnLysIleLeuGlnArgLeuLysThrLeuGluArg 160

Db 421 CAGCACCAAGCTGCTCTCTCAGTATCAGAGATCTTGAAGAGATTCAGAGCGTTAGAGAG 480
 QY 161 GluIleSerGlyGlyAlaMetAlaValAlaValLeuLeuAsnAsnIysLeuTrpVal 180
 Db 481 GAAATTTCCGGAGAGGGCCATGCGCGTTGTGGCGTCTCTTCAACAAGAGCTTACGTC 540
 QY 181 AlaAsnValGlyThrAsnAlaLeuLeuCyLeuSerThrValAspGlyLeuGlnVal 200
 Db 541 GCCAATGCGGTACAAACCGTGACCTTTATGCAAAATCGACATGTGATGGTTGACGGTG 600
 QY 201 ThrGlnLeuAsnValAspHisThrThrGluAsnGluAspGlyLeuPheArgLeuSerGln 220
 Db 601 ACACAGCTGAACGTGGAGCCACACACAGAGAAAGAGATAGCTCTTCCGCTTTGGCAG 660
 QY 221 LeuGlyLeuAspAlaGlyLysIleLysGlnValGlyIleLysGlyGlnIleuSerThr 240
 Db 661 CTGGGCTTGGATGCTGGAAAGATCAAGCAGTGGGATCATCTGTGGGCGAGAGACACC 720
 QY 241 ArgArgIleGlyAspTrpLysValLysTrpGlyTyrThrAspIleAspLeuLeuSerAla 260
 Db 721 CGGGGATCGGGATTACAAAGTTAATAGCTTACACGAGACATTGACCTTCCACGCT 780
 QY 261 AlaLysSerLysProIleIleAlaGluProGluIleHisGlyAlaGlnProLeuAspGly 280
 Db 781 GCCAAGTCCAAACCATCATCGACAGCCGAAATCCATGGGCGACACCGCTGGATGG 840
 QY 281 ValThrGlyPheLeuValLeuMetSerGlyLysLeuTrpLysAlaLeuGluAlaHis 300
 Db 841 GTGACGGGCTTCTTGCTGATGTGGAGGGTTGTAACAAGGCCCTTAGAGGACGCCAT 900
 QY 301 GlyProGlyGlnAlaAsnGlnIleAlaAlaMetIleAspThrGluPheAlaLysGln 320
 Db 901 GGGCTGGGCGAGGCCAACAGGAGATTGCTGCATGTTACACTGATGTTGCCAAGCAG 960
 QY 321 ThrSerLeuAspAlaValAlaGlnAlaValAlaAspArgValLysArgIleHisSerAsp 340
 Db 961 ACCTCCTCGAGCAGAGTGGCCGACGCGCTGCTGAGACGGGTGAAGCCATCCACAGCAG 1020
 QY 341 ThrPheAlaSerGlyGlyGluArgAlaArgPheCysProArgHisGluAspMetThrLeu 360
 Db 1021 ACCTTCGCGACAGTGGTGGGAGCGTGCAGAGTTCTGCCCGCCGACGAGCATGACCTTG 1080
 QY 361 LeuValArgAsnPheGlyTyrProLeuGlyGluMetSerGlnProThrProSerProAla 380
 Db 1081 CTAGTGAAGAACTTTGGCTACCGCTGGGCGAAATGAGCCAGCCACACCGAGCCAGCC 1140
 QY 381 ProAlaIaGlyGlyArgValTyrProValSerValProTrpSerSerAlaGlnSerThr 400
 Db 1141 CCAGCTGACAGAGACAGAGTACCTGTGTCTGTGCAATCTCCAGCGCCGACGAGACCC 1200
 QY 401 SerLysThrSerValThrLeuSerLeuValMetProSerGlnIleGlnMetValaAsnGly 420
 Db 1201 AGCAAGACACAGCGTGAACCTCTCCCTTGTATGACCTCCAGGCGCAGATGTCACAGGG 1260
 QY 421 AlaHisSerAlaSerThrThrLeuAspGluAlaThrProThrLeuThrAsnGlnSerProThr 440
 Db 1261 GCTCACAGTGTCTTCCACCTGAGAGAGCCACCCCACTTCACCAACCAACCCGAGCC 1320
 QY 441 LeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSerSerAspGlyGly 460
 Db 1321 TTAACTCTGCGATCCACCAACAGCAGACAGCAGAGAGCTCCAGCTTGAAGAGGC 1380
 QY 461 LeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGlyArgValGluPro 480
 Db 1381 CTCTTCCGCTCCGGCCCGCCCACTCGCTCCGCGCTGGCAGAGAGCGTGTGTTAGGCC 1440
 QY 481 TyrValAspPheAlaGluLysPheTyrArgLeuTrpSerValAspHisGlyGluGlnSerVal 500
 Db 1441 TATGTGACTTGTGCTGATTTTACCGGCTTGAGAGCTGACCAATGGCGAGCAGAGGCTG 1500
 QY 501 ValThrAlaPro 504


```
QY 321 ThrSerLeuAspAlaValAlaGlnAlaValAspArgValIysArgIleHisSerAsp 340
Db 990 ACCTCCCTGAGACGAGGCGCCAGCGCTGCGAGCCGGGTGAAGGCAATCCACAGCGAC 1049
QY 341 ThrPheAlaSerGlyGlyAlaArgAlaArgPheCysProArgHisGluAspMetThrLeu 360
Db 1050 ACCTTGGCCAGTGGTGGGGAGCGTGCAGGTTCTGCCCCCGGACAGAGACATGACCCCTG 1109
QY 361 LeuValAlaArgAsnPhgIleTyrProLeuGlyGlyMetSerGlnProThrProSerProAla 380
Db 1110 CTAGTACGAGAACTTGTGCTACCCGCTGGGGGAATAGCCAGCCACCGAGCCAGCCAGCC 1169
QY 381 ProAlaAlaGlyGlyArgValTyrProValSerValProTyrSerSerAlaGlnSerThr 400
Db 1170 CCGAGCTGACAGAGAGAGAGTACCTGTGTCTGTCCATCTCCAGCCGCGCCAGAGCACC 1229
QY 401 SerLeuThrSerValThrLeuSerLeuValMetProSerGlnGlyGlnMetValAsnGly 420
Db 1230 AGCAAGACCCAGCGTGAACCTCTCCCTGTTCATGCCCTCCAGGGCCAGATGGTCAACGGG 1289
QY 421 AlaHisSerAlaSerThrLeuAspGluAlaThrProThrLeuThrAsnGlnSerProThr 440
Db 1290 GCTCACAGTGTCTCCACCTCGAGCAAGCCACCCCACTCACCACCAAGCCGAGACC 1349
QY 441 LeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSerAspGlyGly 460
Db 1350 TTAACTCTGACGTCCACCAACGACGACGACGACGACGACGACGACGACGACGACGACG 1409
QY 461 LeuPheArgSerArgProAlaHisSerLeuProProGlyGlyAspGlyArgValGluPro 480
Db 1410 CTCTTCGCTCCCGGCCCGCCCACTCGCTCCCGCTCGGAGAGACGGTGTGTGAAGCCC 1469
QY 481 TyrValAspPheAlaGluPheTyrArgLeuTyrPseValAspHisGlyGlnGlnSerVal 500
Db 1470 TATGTGAGACTTGTGAGTTTATCCGCTCTGGAGCGGTGAGACCATGGCAGACGCGTG 1529
QY 501 ValThrAlaPro 504
Db 1530 GTGACAGCACCG 1541

RESULT 3
AAK56278
ID AAK56278 standard; DNA; 1560 BP.
XX
AC AAK56278;
XX
DT 21-JUL-1999 (first entry)
XX
DE Human TAB1 encoding DNA.
XX
KW Human; TAB1; screening; inhibition; TGF-beta;
KW transforming growth factor beta; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 30..1544
FT tag= a
XX
PN WO9921010-A1.
XX
PD 29-APR-1999.
XX
PF 22-OCT-1998; 98WO-JP04796.
XX
PR 22-OCT-1997; 97JP-0290188.
XX
PA (CHUS ) CHUGAI SEIYAKU KK.
XX
PI Ohmoto T, Ono K, Tsuchiya M;
XX
DR WPI: 1999-312645/26.
DR P-PSDB; AAI09541.
```

```
XX
FT Screening for TGF- beta inhibitory substances, which are useful as
XX drugs for treatment of diseases relating to its disorder
XX
PS Example 1; Page 143-147; 195pp; Japanese.
XX
CC A method has been developed for screening for substances which inhibit
CC the binding of TAK1 polypeptide to TAB1 polypeptide. The method
CC comprises: (a) contacting the polypeptide in the presence of a sample;
CC and (b) detecting the amount of bound polypeptide, in which the sample
CC can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming
CC growth factor (TGF)-beta inhibitory substances can be used in drugs for
CC indications e.g. as TGF-beta signal transduction inhibitors or
CC activators, or extracellular matrix protein production enhancement
CC inhibitors or activators, or cell proliferation prevention inhibitors or
CC activators, or monocyte migration inhibitors or activators, or
CC physiological activity induction inhibitors or activators, or
CC immunosuppression inhibitors or activators, and such substances can also be
CC precipitaton inhibitors or activators, and such substances can also be
CC inhibitors of the TAK1 polypeptide function, particularly kinase
CC activity. The present sequence encodes human TAB1.
XX
SQ Sequence 1560 BP; 332 A; 469 C; 480 G; 279 T; 0 other;

Alignment Scores:
Pred. No.: 8.97e-185 Length: 1560
Score: 2580.00 Matches: 504
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-830-144-4 (1-504) x AAK56278 (1-1560)
QY 1 MetAlaAlaGlnArgArgSerLeuLeuGlnSerGluGlnInProSerTrrThrAspArg 20
Db 30 ATGGCGGCGAGAGAGAGGAGACTTGTCTGAGAGTGAAGCAGACGCAAGCTGGAAGATGAC 89
QY 21 LeuProLeuCyHisLeuSerGlyValGlySerAlaSerAsnArgSerTyrSerAlaAsp 40
Db 90 CTGCCTCTGTGCCACCTCTCTGGGGGTGGCTCAGCTCCCAACCGAGCTACTGTGTGAT 149
QY 41 GlyIysGlyThrGlnSerHisProProGluAspSerTrrLeuLysPheArgSerGluAsn 60
Db 150 GGCAGAGGCACTGAGACCCACCGCAGAGACAGCTGGCTCAAGTTCAGAGATGGAAC 209
QY 61 AsnCyPheLeuTyrGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnPhaVal 80
Db 210 AACTGCTTCTGTATGGGGGCTTTCACAGGCTTATGATGGCAACCGAATGACCACTTCGTG 269
QY 81 AlaGlnArgLeuSerAlaGluLeuLeuGlyGlnLeuAsnAlaGluHisAlaGluAla 100
Db 270 GCCAGCGGCTGTCCGAGAGACTCTGTGGCCAGCTGATGCCAGACGCGCGAGGCC 329
QY 101 AspValArgArgValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGlnUser 120
Db 330 GATGTGGGGGTGTGTGCTGAGCGCTTCGATGTGTGTGAGAGGAGCTTCCCTGGAGTCC 389
QY 121 IleAspAspAlaLeuAlaGluValAspLeuGlnSerGlnLeuProGluGlyValPro 140
Db 390 ATTGACAGCGCTTGGCTGTGAAGGCAAGCTTCAGTTCGCAATTTGCCAGAGGGAGTCCCT 449
QY 141 GlnHisGlnLeuProProGlnTyrGlnLysIleLeuGluArgLeuLysThrLeuGluArg 160
Db 450 CAGACACGAGCTGCTCCTCATGATCAGAAAGATCTTGAAGACATCAAGACGTTAGAGAGG 509
QY 161 GluIleSerGlyGlyAlaMetAlaValAlaValLeuLeuAsnLysLeuTyrVal 180
Db 510 GAAATTTCCGGAGGGGCGCATGGCCGTGTGGCCGCTCTTCTCAACCAAGCTCTACGTC 569
QY 181 AlaAsnValGlyThrAsnArgAlaLeuLeuCyLysSerThrValAspGlyLeuGlnVal 200
Db 570 GCCAATTCGGTACAAACCGTGCATTTTATGCAAAATGCAAGTGGATGGTTCAGCGTG 629
```

```
QY 201 ThrGlnLeuAsnValAspHisThrThrGluAsnGluAspGluLeuPheArgLeuSerGln 220
|
|
|
Db 630 ACACAGCTGAACGTGCACACACACAGAGACGAGATGAGCTTCCGTCTTTCGCAG 689
|
|
|
QY 221 LeuGlyLeuAspAlaGlyIysIleLysGlnValGlyIleCysGlyGlnGluSerThr 240
|
|
|
Db 690 CTGGCTTGGATCTGCGAAAGATCAAGACAGGTGGGATCATCTGTGGCGAGAGACACC 749
|
|
|
QY 241 ArgArgIleGlyAspTyrIysValIysTyrGlyTyrThrAspIleAspLeuSerAla 260
|
|
|
Db 750 CGCGCGATCGGGGATTACAGGTTAAATATGGCTACCGGACATTGACCTTCTCAGCGCT 809
|
|
|
QY 261 AlaLysSerLysProIleIleAlaGluProGluIleHisGlyValaGlnProLeuAspGly 280
|
|
|
Db 810 GCCAAGTCAACAACCATCATCGCAGAGCCAGAAATCATGGGGCACAGCCGTGGATGG 869
|
|
|
QY 281 ValThrGlyPheLeuValLeuMetSerGluGlyLeuTyrIysAlaLeuAlaHis 300
|
|
|
Db 870 GTGACGGCTTCTTGGTGTGCTGATGTCGAGGGGTTGTACAAGGCCCTAGAGCAGCCCAT 929
|
|
|
QY 301 GlyProGlyGlnAlaAsnGlnGluIleAlaAlaMetIleAspThrGluPheAlaLysGln 320
|
|
|
Db 930 GGGCCTGGGCAGGCCAACACAGAGATTGTCGGATGATTCACACTGAGTTTCCCAAGCAG 989
|
|
|
QY 321 ThrSerLeuAspAlaValAlaGlnAlaValAlaAspArgValLysArgIleHisSerAsp 340
|
|
|
Db 990 ACCTCCCTGGACGAGTGGGCCAGGCCGCTGCTGGACCGGGTGAAGCGCATCCACAGCGAC 1049
|
|
|
QY 341 ThrPheAlaSerGlyGlyGluArgAlaArgPheCysProArgHisGluAspMetThrLeu 360
|
|
|
Db 1050 ACCTTCGCCAGTGGTGGGAGCGTGCAGGTTCTGCCCGCCGCGCAGGACATGACCCGTG 1109
|
|
|
QY 361 LeuValArgAsnPheGlyTyrProLeuGlyGluMetSerGlnProThrProSerProAla 380
|
|
|
Db 1110 CTAGTGAGGAACCTTGGCTACCCGCTGGGCGAAATGAGCGAGCCACCGAGCCGAGCC 1169
|
|
|
QY 381 ProAlaAlaGlyGlyArgValTyrProValSerValProTyrSerSerAlaGlnSerThr 400
|
|
|
Db 1170 CCAGCTGCAGGAGCAGAGTGATACCCCTGTGTCTGTGCCATATCTCCAGCGCCAGCACC 1229
|
|
|
QY 401 SerLysThrSerValThrLeuSerLeuValMetProSerGlnGlyGlnMetValAsnGly 420
|
|
|
Db 1230 AGCAAGACCAACGCGTGACCCCTCTCCCTTGTCTATGCCCTCCCGAGGCCAGATGGTCAACGGG 1289
|
|
|
QY 421 AlaHisSerAlaSerThrLeuAspGluAlaThrProThrLeuThrAsnGlnSerProThr 440
|
|
|
Db 1290 GCTCACAGTGTCTCCACCTGGACGAGCCACCCCTCCACCAACCAAGCCCGACCC 1349
|
|
|
QY 441 LeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSerSerSerGlyGly 460
|
|
|
Db 1350 TTAACCTTCAGTCCACCAACACGACACGACGACGACGAGCTCCAGCTTCGACGGAGGC 1409
|
|
|
QY 461 LeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGlyArgValGluPro 480
|
|
|
Db 1410 CTCTTCGCTCCCGGCCGCCACCTCGCTCCCGCTGGCGAGACGCTGCTGTGAGGCC 1469
|
|
|
QY 481 TyrValAspPheAlaGluPheTyrArgLeuThrSerValAspHisGlyGluGlnSerVal 500
|
|
|
Db 1470 TATGTGGACTTTGTGAGTTTATCCGCTCTGGAGCGTGGACCATGGCGAGCAGCGGTG 1529
|
|
|
QY 501 ValThrAlaPro 504
|
|
|
Db 1530 GTACAGCACCG 1541
|
|
|
RESULT 4
AAA39106
ID AAA39106 standard; DNA: 1560 BP.
XX
AC AAA39106;
XX
DT 04-SEP-2000 (first entry)
XX
```

```
DE Human TAB-1 nucleotide sequence SEQ ID NO:3.
XX
KW Human; TAK-1; TAB-1; mitogen activated protein kinase; MAPK;
KW screening; signal transduction; inhibition; inflammatory cytokine;
KW IL-1; interleukin 1; TNF; tumour necrosis factor; inflammation;
KW antiinflammatory; suppression; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 30..1544
FT /tag= a
FT /product= "TAB-1"
XX
FN W0200023610-A1.
XX
PD 27-APR-2000.
XX
XX 21-OCT-1999; 99WO-JP05817.
XX
XX 21-OCT-1998; 98JP-0299962.
XX
XX (CHUS ) CHUGAI SEIYAKU KK.
XX
XX Tsuchiya M, Ohtomo T, Sugamata Y, Matsumoto K;
XX WPI; 2000-339707/29.
XX P-PSDB; AAY91001.
XX
PT Method for screening inhibitors of TAK1 signal transduction for
PT suppression of inflammatory cytokine production and use as
PT antiinflammatory agents -
XX
XX Disclosure; Page 85-90; 100pp; Japanese.
XX
XX The present invention describes a method for screening compounds for
XX inhibition of inflammatory cytokine signal transduction by contacting
XX the sample with TAK1 and its receptor TAB1 and selecting for inhibition
XX of TAK1/TAB1 binding. Also described is a method for screening compounds
XX for inhibition of inflammatory cytokine signal transduction in which the
XX inhibition of TAK1 phosphorylation is selected for; and drug
XX compositions for the treatment of inflammatory disorders containing as
XX active component an inflammatory cytokine signal transduction inhibitor.
XX TAK1 is an essential component of the signalling process which results
XX in release of inflammatory cytokines such as interleukin-1 (IL-1),
XX IL-10, tumour necrosis factor (TNF) and IL-6. The methods can be used
XX for the selection of effective antiinflammatory agents. The present
XX sequence encodes human TAB-1, which is used in the exemplification of the
XX present invention.
XX
SQ Sequence 1560 BP; 332 A; 469 C; 480 G; 279 T; 0 other;

Alignment Scores:
Pred. No.: 8.97e-185 Length: 1560
Score: 2580.00 Matches: 504
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 21 Gaps: 0

US-09-830-144-4 (1-504) x AAA39106 (1-1560)
QY 1 MetAlaAlaGlnArgArgSerLeuLeuGlnSerGluGlnGlnProSerTrpThrAspAsp 20
|
|
|
Db 30 ATGGCGCGCGCAGAGGAGGAGCTTCTGCAGAGTGCAGCAGCAGCCAGCTGCAGCAGTGAC 89
|
|
|
QY 21 LeuProLeuCysHisLeuSerGlyValGlySerAlaSerAsnArgSerTyrSerAlaAsp 40
|
|
|
Db 90 CTGCTCTCTGCCACCTCTCTGGGGTTGGCTCAGCTCCAAACCCAGCTACTCTGCTGAT 149
|
|
|
QY 41 GlyLysGlyThrGluSerHisProGluAspSerTrpLeuLysPheArgSerGluAsn 60
|
|
|
Db 150 GGCAAGGGCACTGAGAGCCACCCGCGAGAGGACAGCTGGCTCAAGTTCCAGGAGTGAGAAC 209
|
|
|
```

```

QY 61 AsnCysPheLeuTyrGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnPheVal 80
DB 210 AACTGCTTCTGTATGGGGTCTTCAACGGCTATGATGGCAACCGAGTGACCAACTTCGTGG 269
QY 81 AlaGlnArgLeuSerAlaGlnLeuLeuLeuGlnGlnLeuAsnAlaGlnHisAlaGlnVal 100
DB 270 GCCCAGCGGCTGTCCGACAGACTCTCTCTGGGCGACCTAAATGCGAGCAACCCGAGGCC 329
QY 101 AspValArgArgValLeuLeuGlnAlaPheAspValAlaGlnArgSerPheLeuGlnSer 120
DB 330 GATGTGCGGGGTGTGCTGTGAGAGGCTTCGATGTGTGGAGAGGAGCTTCCTGGAGTCC 389
QY 121 IleAspAspAlaLeuAlaGlnLysAlaSerLeuGlnSerGlnLeuProGlnGlyValPro 140
DB 390 ATTGAAGAGCGCTTGGCTGAGAGGCAACCTCCCAATTCGCAAGGAGAGTCTCCCT 449
QY 141 GlnHisGlnLeuProProGlnTyrGlnLysIleLeuGlnArgLeuLysThrLeuGlnArg 160
DB 450 CAGCACCAGCTGCTCTCTCAAGTATCCAGAAAGATCTTGAGAGACTCAAGACGTTAAGAGG 509
QY 161 GlnIleSerGlyValAlaMetAlaValAlaValLeuLeuAsnAsnLysLeuTyrVal 180
DB 510 GAAATTTCCGGAAGGGCCATGGCCGTGTGGCGTCTTCTCAACAACAAGCTTACGTC 569
QY 181 AlaAsnValGlyThrAsnArgAlaLeuLeuCysLysSerThrValAspGlyLeuGlnVal 200
DB 570 GCCAATGTGCTGATCAAAACCGTGCACCTTTATGCAATCCACAGATGATGGTTCAGCGTG 629
QY 201 ThrGlnLeuAsnValAspHisThrThrGlnAsnGlnAspGlnLeuPheArgLeuSerGln 220
DB 630 AACACCTTAACGCTGACCAACACACAGAAACAGAGATGACCTCTTCGTTTCGACG 689
QY 221 LeuGlyLeuAspAlaGlyLysIleLysGlnValGlyIleIleCysGlyGlnLysSerThr 240
DB 690 CTGGGCTTGATGCTGAAAGATCAAGCAGGTGGGATCATCTGTGGAGAGAGCAACC 749
QY 241 ArgArgIleGlyAspTyrLysValLysTyrGlyTyrThrAspIleAspLeuLeuSerAla 260
DB 750 CGGCGGATCGGGGATTAAGAAGTTAAATGTGCTACACGACATTTGACCTTCTCCGCGCT 809
QY 261 AlaLysSerLysProIleIleAlaGlnProGlnIleHisGlyAlaGlnProLeuAspGly 280
DB 810 GCCAATCCAAACCAATCATCCGACAGCCAGAAATCCATGGGCGACAGCCGTGATGGG 869
QY 281 ValThrGlyPheLeuValLeuMetSerGlnGlyLeuTyrLysAlaLeuGlnAlaHis 300
DB 870 GTGACGGGCTCTTGGTGTGATGTGAGGGGTTGTACAAGCCCTAGAGGACGCCAT 929
QY 301 GlyProGlyGlnAlaAsnGlnLysIleAlaAlaMetIleAspThrGlnPheAlaLysGln 320
DB 930 GGGCGCTGGGAGCCAAACAGAGATGTCTGCAATGATGACACGATTTGGCCAAAGAG 989
QY 321 ThrSerLeuAspAlaValAlaGlnAlaValAlaAspArgValLysArgIleHisSerAsp 340
DB 990 AACTCTCTTGAAGCAGTGGCCCAAGCCGTCTGACCGGGGTAAACGCATCCACAGCCATC 1049
QY 341 ThrPheAlaSerGlyGlyGlnArgAlaArgPheCysProArgHisGlyAspMetThrLeu 360
DB 1050 ACCTTCGCGAGTGTGGGAGCGTGCAGGTTCTGCCCCCGGACAGAGACATGACCTGTG 1109
QY 361 LeuValArgAsnPheGlyTyrProLeuGlyGlnMetSerGlnProThrProSerProAla 380
DB 1110 CTAAGTGAGAACTTTGGCTACCCGCTGGGCGAAATGAGCACCCACCGAGCCACCC 1169
QY 381 ProAlaAlaGlyArgValTyrProValIserValProTyrSerSerAlaGlnSerThr 400
DB 1170 CCAAGCTCAGGAGAGAGAGTGAACCTGTGTGTGCACTACTCCAGGCGCCAGAGCACCC 1229
QY 401 SerLysThrSerValThrLeuSerLeuValMetProSerGlnGlyGlnMetValAsnGly 420
DB 1230 ACCAAGACACAGCGTACCTCTCTGTGCAATGCCCTCCAGAGGCGCAATGTCACAGCGG 1289

```

```

QY 421 AlaHisSerAlaSerThrLeuAspGlnAlaThrProThrIleuThrAsnGlnSerProThr 440
DB 1290 GCTCACAGTGTCTTCCACCTCGGAGCAAGCCACCCCACTTCACCAACCAACCCGAGCC 1349
QY 441 LeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSerAspGlyGly 460
DB 1350 TTAACTTCGAGTCCACCAACGACACGACAGAGAGAGCTTCAGCTTCAGCGAGGC 1409
QY 461 LeuPheArgSerArgProAlaHisSerLeuProProGlyGlnAspGlyArgValGlnPro 480
DB 1410 CTTTCCTCGCTCCGCGCCGCCCACTGCTCCGCTGGGAGAGACGGTGTGTAGGCC 1469
QY 481 TyrValAspPheAlaGlnPheTyrArgLeuTyrPserValAspHisGlyGlnGlnSerVal 500
DB 1470 TATGTGACTTTGTGTGATTTACCGCTTCGAGCGTGGACCATGCGAGCAGACGCTG 1529
QY 501 ValThrAlaPro 504
DB 1530 GTGACAGCACCG 1541

RESULT 5
ID AAX56282
XX AAX56282 standard; DNA; 1569 BP.
AC AAX56282;
XX
XX 21-JUL-1999 (first entry)
DT
XX
DE Human TAB1-FLAG encoding DNA.
XX
XX Human; TAB1; TAK1; screening; inhibition; TGF-beta;
XX transforming growth factor beta; ss.
XX
XX Homo sapiens.
OS
XX Synthetic.
OS
XX
FH Key Location/Qualifiers
FT CDS 7..1560
FT tag= a
XX
XX MO9921010-At.
XX
XX 29-APR-1999.
XX
XX 22-OCT-1998; 98WO-JP04796.
XX
XX 22-OCT-1997; 97JP-0290188.
XX
XX (CHUS ) CHUGAI SEIYAKU KK.
XX
XX Ohtomo T, Ono K, Tsuchiya M;
XX
XX WPI; 1999-312645/26.
XX
XX P-PSDB; AAY09546.
XX
XX
PT Screening for TGF- beta inhibitory substances, which are useful as
PT drugs for treatment of diseases relating to its disorder
XX
XX Example 1; Page 159-163; 1995pp; Japanese.
XX
XX A method has been developed for screening for substances which inhibit
XX the binding of TAK1 polypeptide to TAB1 polypeptide. The method
XX comprises: (a) contacting the polypeptide in the presence of a sample;
XX and (b) detecting the amount of bound polypeptide, in which the sample
XX can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming
XX growth factor (TGF)-beta inhibitory substances can be used in drugs for
XX indications e.g. as TGF-beta signal transduction inhibitors or
XX activators, or extracellular matrix protein production enhancement
XX inhibitors, or activators, or cell proliferation prevention inhibitors or
XX activators, or monocyte migration inhibitors or activators, or
XX physiological activity induction inhibitors or activators, or
XX immunosuppression inhibitors or activators, or amyloid beta protein
XX precipitation inhibitors or activators, and such substances can also be

```

CC inhibitors of the TAK1 polypeptide function, particularly kinase
CC activity. The present sequence encodes TAB1-FLAG from an example of
CC the present invention.

XX Sequence 1569 BP; 343 A; 466 C; 476 G; 284 T; 0 other;

Alignment Scores:

Pred. No.: 9,04e-185 Length: 1569
Score: 2580.00 Matches: 504
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 20 Gaps: 0

US-09-830-144-4 (1-504) x AAX56282 (1-1569)

QY 1 MetAlaAlaGlnArgSerLeuLeuGlnSerGluGlnProSerTrpThrAspAsp 20
Db 7 ATGGCGGCGCAGAGGAGGCTTCTGCAGAGTGAGCAGCAAGCTGGACAGATGAC 66
QY 21 LeuProLeuCysHisLeuSerGlyValGlySerAlaSerAsnArgSerTyrSerAlaAsp 40
Db 67 CTGCCTCTCTGCCACCTCTCTGGGTTGGCTCAGCCTCCAAACCGCAGCTACTCTGCTGAT 126
QY 41 GlyLysGlyThrGluSerHisProProGluAspSerTrpLeuLysPheArgSerGluAsn 60
Db 127 GGCAAGGGCACTGAGAGCCACCCGCCAGAGACAGCTGGCTCAAGTTTCAGAGTGAGAAC 186
QY 61 AsnCysPheLeuTyrGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnPheVal 80
Db 187 AACTGCCTTCCTGTATGGGCTTCCACGGCTATGATGGCAACCGAGTGACCAACTTCGTG 246
QY 81 AlaGlnArgLeuSerAlaGluLeuLeuGlyGlnLeuAsnAlaGluHisAlaGluAla 100
Db 247 GCCCAGCGGCTGTCGAGAGCTCTCTGGCGAGCTGAATGCCAGACGCCGAGGCC 306
QY 101 AspValArgArgValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGluSer 120
Db 307 GATGTGGCGGTGTGCTGCTGCAGGCCTTCGATGTGTGTGGAGGAGCTTCCTGGAGTCC 366
QY 121 IleAspAlaLeuAlaGluLysAlaSerLeuGlnSerGlnLeuProGluGlyValPro 140
Db 367 ATTGACACGCTTGGCTGAGAGGCAAGCCTCCAGTCGCAATGGCAGAGGAGTCCCT 426
QY 141 GlnHisGlnLeuProGlnTyrGlnLysIleLeuGluArgLeuLysThrLeuGluArg 160
Db 427 CAGCACCGCTGCTCTCTAGTATCAGAGATCTCTTGAGAGACTCAAGAGCTTAGAGAG 486
QY 161 GluIleSerGlyAlaMetAlaValAlaValLeuLeuAsnAsnLysLeuTyrVal 180
Db 487 GAAATTTTCGGAGGGGCATGGCGTTGTGGCGTCTCTCAACAACAAGCTCTACGTC 546
QY 181 AlaAsnValGlyThrAsnArgAlaLeuLeuCysLysSerThrValAspGlyLeuGlnVal 200
Db 547 GCCAATGTCGGTACAAACCGTGACCTTTTATGCAAAATCGACAGTGGATGGTTGCAGGTG 606
QY 201 ThrGlnLeuAsnValAspHisThrThrGluAsnGluAspGluLeuPheArgLeuSerGln 220
Db 607 ACACAGCTGAACGTGGACCAACACACAGAGAGATGAGCTCTTCGCTCTTCGCGAG 666
QY 221 LeuGlyLeuAspAlaGlyLysIleLysGlnValGlyIleIleCysGlyGlnGluSerThr 240
Db 667 CTGGCTTGGATGCTGGAAGATCAAGCAGGTGGGATCATCTGTGGCAGGAGACACC 726
QY 241 ArgArgIleGlyAspTyrLysValLysTyrGlyTyrThrAspIleAspLeuLeuSerAla 260
Db 727 CGCGGATCGGGATTACAAAGGTAAATATGCTGTACACGACATTGACCTTCTCAGCGCT 786
QY 261 AlaLysSerLysProIleIleAlaGluProGluIleHisGlyAlaGlnProLeuAspGly 280
Db 787 GCCAAGTCCAAACAATCATTCGAGAGCCAGAAATCATGGGGCAGCGCTGGATGGG 846
QY 281 ValThrGlyPheLeuValLeuMetSerGluGlyLeuTyrLysAlaLeuGluAlaHis 300

Db 847 GTGACGGGCTTCTTGCTGCTGATGTCGAGGGGTGTATCAAGGCCCTAGAGCGACCCAT 906
QY 301 GlyProGlyGlnAlaAsnGlnGluIleAlaMetIleAspThrGluPheAlaLysGln 320
Db 907 GGGCCTGGGCAGGCCAACCCAGAGATTGCTGCGATTGACACTGAGTTTGTCCCAAGCAG 966
QY 321 ThrSerLeuAspAlaValAlaGlnAlaValValAspArgValLysArgIleHisSerAsp 340
Db 967 ACCTCTCTGGACCGAGTGGCCAGGCCCTGCTGGACCGGGTGAAGCGCATCCACAGCGAC 1026
QY 341 ThrPheAlaSerGlyGlyGluArgAlaArgPheCysProArgHisGluAspMetThrLeu 360
Db 1027 ACCTTCCGACAGTGTGGGAGCGTGCCAGGTTCTGCCCCCGCAGCAGACATGACCTG 1086
QY 361 LeuValArgAsnPheGlyTyrProLeuGlyGluMetSerGlnProThrProSerProAla 380
Db 1087 CTAGTGAGGAACTTTGGCTACCCGCTGGGCGAAATGAGCCAGCCACACCGAGCCAGCC 1146
QY 381 ProAlaAlaGlyGlyArgValTyrProValSerValProTyrSerSerAlaGlnSerThr 400
Db 1147 CCAGCTGCGAGGAGCAGGTGATCCCTGTGTCTGTGCCATTACTCCAGCGCCAGAGCACC 1206
QY 401 SerLysThrSerValThrLeuSerLeuValMetProSerGlnGlyGlnMetValAsnGly 420
Db 1207 AGCAAGACCCAGCGTGACCTCTCCCTTGTCTCATGCCCTCCAGGGCCAGATGTTCAACGG 1266
QY 421 AlaHisSerAlaSerThrLeuAspGluAlaThrProThrLeuThrAsnGlnSerProThr 440
Db 1267 GCTCAGTGTCTTCACCTCGAGCAGCACCACCCCTCAGCAACCAAGGCCGACC 1326
QY 441 LeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSerSerSerAspGly 460
Db 1327 TTAACCTTGCACTCCACCAACACGACACGACAGCAGCAGCTCCAGCTCTGACGGAGGC 1386
QY 461 LeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGlyArgValGluPro 480
Db 1387 CTCCTCCGCTCCCGGCCGCCCACTCGCTCCCGCTGGCAGGAGCGTCTGTTGAGCCC 1446
QY 481 TyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGlyGluGlnSerVal 500
Db 1447 TATGTGGACTTTGTGTAGTTTTACCGCCTCTGGAGCGTGGACCATGGCAGCAGAGCGTG 1506
QY 501 ValThrAlaPro 504
Db 1507 GTGACAGCACCG 1518
RESULT 6
AAX56310
ID AAX56310 standard; DNA; 1568 BP.
XX AAX56310;
AC AAX56310;
XX 21-JUL-1999 (first entry)
DT Human TAB1 encoding DNA SEQ ID NO:42.
DE Human; TAB1; TAK1; screening; inhibition; TGF-beta;
KW transforming growth factor beta; ss.
XX Homo sapiens.
FH Key Location/Qualifiers
FT CDS 11..1552
FT /*tag= a
XX
PN WO9921010-A1.
XX
PD 29-APR-1999.
XX
PF 22-OCT-1998; 98WO-JP04796.
XX
PR 22-OCT-1997; 97JP-0290188.

AAT91178
ID AAT91178 standard; cDNA; 1560 BP.
AC AAT91178;
XX 14-APR-1998 (first entry)
DT Human TAB1 (TAK1 binding protein) cDNA.
XX
XX TAB1; TAK1 binding protein; transforming growth factor-beta;
KW signal transduction; human; ds.
KW Homo sapiens.
OS
XX
FH Key Location/Qualifiers
CDS 30..1544
FT /*tag= a
FT variation 185
FT /*tag= b
FT /note= "another clone has cytosine at position 185,
FT with codon AGA (Arg) altered to AGC (Ser)"
XX
PN EP803571-A2.
XX
PD 29-OCT-1997.
XX
PF 24-APR-1997; 97EP-0302808.
XX
XX 20-NOV-1996; 96US-0752891.
PR 24-APR-1996; 96JP-0126282.
PR 28-OCT-1996; 96JP-0300856.
XX
PA (UENO/) UENO N.
XX
XX Matsumoto K, Nishida E;
XX WPI; 1997-515318/48.
DR P-PSDB; AAW26707.
XX
XX DNA encoding TAK1 binding protein TAB1 - member of transforming
PT growth factor beta receptor signal production pathway, which
PT activates TAK-1 kinase activity upon binding
PT
XX Example 5; Page 19-21; 30pp; English.
XX
XX This cDNA clone codes for human TAB1 (see AAW26707), a novel member
CC of the transforming growth factor-beta receptor signal transduction
CC pathway, which activates TAK-1 kinase activity upon binding. To
CC obtain the full-length TAB1 coding sequence, a human kidney cDNA
CC library was screened using as a probe a partial TAB1 cDNA obtained
CC from a yeast two-hybrid assay for proteins that interacted with
CC TAK1. The 5' terminus was identified by 5'RACE. 2 Different
CC clones were sequenced, with cytosine (see AAT91175) and adenine as
CC the 185th nucleotide, respectively, and deposited as FERM BP-5599
CC and FERM BP-5508, respectively. Also claimed are: isolated DNA
CC encoding a protein modified by a substitution, deletion and/or
CC addition of 1 or more amino acids of the 504-residue TAB1 sequence;
CC (2) DNA which can hybridise with the 1560 bp nucleic acid sequence;
CC (3) isolated DNA encoding a protein comprising amino acids 21-579
CC or 437-504 of the 504 TAB1 sequence; (4) DNA encoding a fusion
CC protein comprising an above protein or polypeptide; (5) expression
CC vector comprising an above DNA; and (6) host cell, preferably a
CC mammalian or yeast cell, transformed by the expression vector.
CC Cells expressing TAB1 and TAK1 can be used to screen for TGF-beta
CC signalling pathway inhibitors by contacting the cells with a test
CC compound, and measuring the TAK1 kinase activity.
XX
SQ Sequence 1560 BP; 333 A; 468 C; 480 G; 279 T; 0 other;

Alignment Scores:

Pred. No.: 2,13e-184 Length: 1560
Score: 2575.00 Matches: 503
Percent Similarity: 99.80% Conservative: 0

Best Local Similarity: 99.80% Mismatches: 1
Query Match: 99.81% Indels: 0
DB: 18 Gaps: 0

US-09-830-144-4 (1-504) x AAT91178 (1-1560)

QY 1 MetAlaAlaGlnArgSerLeuLeuGlnSerGluGlnGlnProSerTrpThrAspAsp 20
Db 30 ATGGCGGCGCAGAGGAGCTTGTGCAGAGTGAGCAGCAGCAAGCTGCAGAGTGAC 89
QY 21 LeuProLeuCysHisLeuSerGlyValGlySerAlaSerAsnArgSerTyrSerAlaAsp 40
Db 90 CTGCCTCTCTGCCACCTCTCTGGGGTTGGCTCAGCCTCCAAACCGCAGCTACTCTGCTGAT 149
QY 41 GlyLysGlyThrGluSerHisProGluAspSerTrpLeuLysPheArgSerGluAsn 60
Db 150 GGCAGGGCATGTGAGCACCACCCGCGAGGACAGATGGCTCAAGTTCAGAGGTGAGAAC 209
QY 61 AsnCysPheLeuTyrGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnPheVal 80
Db 210 AACTGCTTCTGTATGGGGTCTTCAACGGCTATGATGCAACCGAGTGACCAACTTCGTG 269
QY 81 AlaGlnArgLeuSerAlaGluLeuLeuLeuGlyGlnLeuAsnAlaGluHisAlaGluAla 100
Db 270 GCCCAGCGGCTGTCCGACAGAGCTCTCTGGGCCAGCTGAATGCCGAGCAGCGCGAGGCC 329
QY 101 AspValArgArgValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGluSer 120
Db 330 GATGTGCGGCGTGTCTGCTGCAGGCCCTTTCGATGTGTGGTGAGAGGAGCTTCTCGAGTCC 389
QY 121 IleAspAspAlaLeuAlaGluLysAlaSerLeuGlnSerGlnLeuProGluGlyValPro 140
Db 390 ATTGACGACGCTTGGCTGAGAGCGCAAGCCTCCAGTCGCAATTCGCCAGGAGGAGTCCCT 449
QY 141 GlnHisGlnLeuProProGlnTyrGlnLysIleLeuGluArgLeuLysThrLeuGluArg 160
Db 450 CAGCACCAGCTGCCTCTCAGTATCAGAAGATCCTTTCAGAGACTCAAGACGTTAGAGAGG 509
QY 161 GluIleSerGlyGlyAlaMetAlaValAlaValAlaValLeuLeuAsnLeuLysLeuTyrVal 180
Db 510 GAATTTCCGGAGGGGCGCATGGCCGTTGTGGCGGTCTCTTCAACAACAAGCTCTACGTC 569
QY 181 AlaAsnValGlyThrAsnArgAlaLeuLeuCysLysSerThrValAspGlyLeuGlnVal 200
Db 570 GCCAATGTCGGTACAAACCGTGCACCTTTATGCAAAATCGACAGTGGATGGTTCAGGTG 629
QY 201 ThrGlnLeuAsnValAspHisThrThrGluAsnGluAspGluLeuPheArgLeuSerGln 220
Db 630 ACACAGCTGAACGTGGACCCACACACAGAGAACGAGGATGAGCTCTTCCGCTTTTCGAG 689
QY 221 LeuGlyLeuAspAlaGlyLysIleLysGlnValGlyIlelleCysGlyGlnGluSerThr 240
Db 690 CTGGCTTGGATGCTGGAAGATCAAGCAGGTGGGGATCATCTGTGGCGAGGAGAGCACC 749
QY 241 ArgArgIleGlyAspTyrLysValLysValLysTyrGlyTyrThrAspIleAspLeuLeuSerAla 260
Db 750 CGCGGATCGGGATTAACAAGTTAATAATGGCTACACGACATGACCTTCTCAGCGCT 809
QY 261 AlaLysSerLysProIlelleAlaGluProGluIleHisGlyAlaGlnProLeuAspGly 280
Db 810 GCCAAGTCCAAACCAATCATCGCAGAGCCAGAAATCCATGGGGCAGACGCGCTGGATGGG 869
QY 281 ValThrGlyPheLeuValLeuMetSerGluGlyLeuTyrLysAlaLeuGluAlaAlaHis 300
Db 870 GTGACGGGCTTCTTGGTGCTGATCTCGAGGGGTGTACAAGGCCCTTAGAGGAGCGCCAT 929
QY 301 GlyProGlyGlnAlaAsnGlnGluIleAlaAlaMetIleAspThrGluPheAlaLysGln 320
Db 930 GGGCTGGGCGAGGCCAACAGGAGATTGCTGCGATGATTCACACTGAGTTTGCCACAGCAG 989
QY 321 ThrSerLeuAspAlaValAlaGlnAlaValValAspArgValLysArgIleHisSerAsp 340
Db 990 ACCTCCCTGGACGAGTGGGCCAGGCCGCTCTGTGACCGCGGTGAAGCGCATCCACAGCGAC 1049

Qy 341 ThrPheAlaSerGlyGlyGluArgAlaArgPheCysProAlaHisGluAspMetThrLeu 360
 Db 1050 ACCTTCGCAGTGTGGGGAGCGTCCAGGTTCTGCCCCCGGACAGAGACATGACCTCG 1109
 Qy 361 LeuValArgAspPheLeuTyrProLeuGlyGlyMetSerGlnProThrProSerProAla 380
 Db 1110 CTAGTGAGGAACCTTGGCTACCCGCTGGCGGAATGAGCCAGCCACACGAGCCAGCC 1169
 Qy 381 ProAlaAlaGlyGlyValTyrProValSerValProTyrSerSerValGlnSerThr 400
 Db 1170 CCAGCTGCAGGAGGAGAGAGTACCTGTGTCTGTGTCATCTCCAGCGCCACAGACACC 1229
 Qy 401 SerLeuThrSerValThrLeuSerLeuValMetProSerGlnGlyGlnMetValSengly 420
 Db 1230 AGCAAGACGAGCGTGAACCTCTCCCTTGTCAATGCCCCCAGGGCCAGTGTCAACGGG 1289
 Qy 421 AlaHisSerAlaSerThrLeuAspGlyAlaThrProThrLeuThrAsnGlnSerProThr 440
 Db 1290 GCTCAGAGTGTCTCAACCTTGAGAGAGCCACCCCACTCCACCAAGCCCGAC 1349
 Qy 441 LeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSerSerSerValGly 460
 Db 1350 TTAACTTCGAGTCCACCAACAGCAGCAGCAGCAGCAGCAGCTCCAGCTGACGAGG 1409
 Qy 461 LeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGlyValArgValGluPro 480
 Db 1410 CTCTTCGCGTCCCGCGCCGCTCGCTCCGCTGGGAGACCGTGTGTGAGCC 1469
 Qy 481 TyrValAspPheAlaGluPheTyrArgLeuTyrSerValAspHisGlyGlnGlnSerVal 500
 Db 1470 TATGTGACTTGTCTGATGTTTACCGCTCTGAGCGTGACCATGCGAGAGAGCTG 1529
 Qy 501 ValThrAlaPro 504
 Db 1530 GTGACGACCG 1541
 RESULT 8
 ID ABA20494 standard; DNA; 16877 BP.
 XX
 AC ABA20494;
 XX
 DT 23-JAN-2002 (first entry)
 XX
 DE Human nervous system related polynucleotide SEQ ID NO 12825.
 XX
 KW Human; nocotropic; neuroprotective; cytoprotective; anti-inflamatory; anti-HIV; antibacterial; virocidic;
 KW immunosuppressive; antistickling; antianaemic; antiathrictic; cancer;
 KW antiParkinsonian; hepatotoxic; cerebroprotective; antiinflammatory;
 KW antidiabetic; antidiabetic; antidiabetic; anticonvulsant; antifungal;
 KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; nephrotoxic; gene therapy; vaccine; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200159063-A2.
 XX
 PD 16-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01334.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189674.
 PR 17-MAR-2000; 2000US-0190676.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 25-SEP-2000; 2000US-0234999.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.

PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 20-OCT-2000; 2000US-0242221.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246509.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0251930.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-541565/60.
XX
PT Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system
PT cancers and metastases -
XX
XX Disclosure; SEQ ID NO 12825; 1701pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (ABA11004-ABA21534) and proteins
CC (AB14678-AB18001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC

CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 16877 BP; 3563 A; 4508 C; 4516 G; 4290 T; 0 other;

Alignment Scores:
Pred. No.: 2,27e-30 Length: 16877
Score: 543.50 Matches: 295
Percent Similarity: 18.86% Conservative: 43
Best Local Similarity: 16.46% Mismatches: 107
Query Match: 21.07% Indels: 1348
DB: 22 Gaps: 19

US-09-830-144-4 (1-504) x ABA20494 (1-16877)

QY 12 GluGlnGlnProSerTrpThrAspAspLeuProLeuCysHisLeuSerGlyValGlySer 31
DB 1 GAGCAGCAGCAAGCTGGACAGATGACCTCTCTCTGCCACCTCTCTGGGGTTGGCTCA 60

QY 32 AlaSerAsnArgSerTyrSerAlaAspGlyLysGlyThrGluSerHisProGluAsp 51
DB 61 GCCTCAACCGCAGCTACTCTGCTGATGGCAAGGCACTGAGAGCCACCCGCCAGGAC 120

QY 52 SerTrpLeuLysPhe----- 56
DB 121 AGCTGGCTCAAGTTCAAGTGTGTGCCAGCATTTCTGTGTGGCCCGGGAGTTGGTT 180

QY 56 ----- 56
DB 181 GGTTTTGAAGCAAGAAAGACACCCACCTTCGAGCTTTCTCGTATGGCTTGGCAGTGAC 240

QY 56 ----- 56
DB 241 ATGTGGCCCGTGAGAGTGGGCTCTGCTGCTGTCTTGCAAGGGCCCTGCTGTATGGGGT 300

QY 56 ----- 56
DB 301 AGCGTGAGCATGGGAGGAGGGTATCCAGAAATGTCATAGCCAGAGTAAATGATGGCTA 360

QY 56 ----- 56
DB 361 AAGCAGGGGACCCAGAGAGGGCCCTGAAAGCTGCAGCTGCTGCTGCTTAGCTCCCCCA 420

QY 56 ----- 56
DB 421 ATTCCTTTCCCTTTCTCCCTCCACCTCCGTCGAGACCCCTGGTCTCAGGCCCTCCCTGTGCC 480

QY 57 -----ArgSerGluAsnAsnCysPheLeuTyrGlyValPheAsnGlyTyrAsp 72
DB 481 TCTCCCTCTTCAGAGGTGAGAACAACTGCTTCCTGTATGGGGCTTTCACCGGTATGAT 540

QY 73 GlyAsnArgValThrAsnPheValAlaGlnArgLeuSerAlaGluLeuLeuGlyGln 92
DB 541 GGCAACCGAGTGACCAACTTGTGGCCCGAGGGCTGTCCGAGAGCTCTCTGCTGGCCAG 600

QY 93 LeuAsnAlaGluHisAlaGluAlaAspValArgArgValLeuLeuGln----- 108
DB 601 CTGAATGCCGAGCAGCGCCGAGGCGCATGTGCGGGTGTGCTGTGCTGCAGGTAATGTCGCG 660

QY 108 ----- 108
DB 661 GGGCCACAGTGACCCAGCCACATCATGTCCCCCACCACCAAGGCTTGGGCCCTGCACCTC 720

QY 108 ----- 108

Db 721 TAGCATGTTGCCAGGGTGTGTGAAGATCTGCGGCCCCCTTCAACCCAGTAGAGAGC 780
Qy 108 ----- 108
Db 781 AGCTCCAGCGTAGAGCCCCCCCCCAACAGAGAGTCCAGGACCAGCAGCAGCAGG 840
Qy 108 ----- 108
Db 841 CCTTCAATGCGTCGACAGAGAGTACTGCTGAGACGCGGGGATTTAGGATGGAGCTT 900
Qy 108 ----- 108
Db 901 GGAGAGAGTGTGAGGTGGAGCAGGGCAAGGCTGTGTAGAAATGGGTCATTTAGAGT 960
Qy 108 ----- 108
Db 961 ACCCTTTCTTCTCTATGTGTGTCAGGTGCTCAGCTCAGGTGACAGAGCATTCCGTGG 1020
Qy 108 ----- 108
Db 1021 CGTCCAGAGAGACCTTGCCCTCCCTTCTGAGGGGCCGCGCCCTCATTTGACTGTCC 1080
Qy 108 ----- 108
Db 1081 ACCTTTCTGCTTGCGCTGTGGCAGTTTCTCTGTGGCATCTGCTTTCAAGACATGTCT 1140
Qy 108 ----- 108
Db 1141 CAGGCCATTTGAGATGAGAGATGGGCTTCTGTTCCCGAGAGGGGGTGTGCAGCCTT 1200
Qy 108 ----- 108
Db 1201 TTTCCTCCCTTACGACCTCAGGCTTCATGCGCATGATTCTCAGCAGATCTCAGACAG 1260
Qy 108 ----- 108
Db 1261 GGGAGAAAGTGTACGTGTCAACAGTGTCTTGCGGCTGTGGGTTTGACAGAAAGCTCCC 1320
Qy 108 ----- 108
Db 1321 AGAAGAAGTTGTAGTAATCATATCAGAGTGTCTATAGCCTGGCAGTTTCTGAGAGATT 1380
Qy 108 ----- 108
Db 1381 GCCTTGCTGAGTGCAGGTTGACTGTCTTACTTAAACCCCTGGAAACACATTAACTTCC 1440
Qy 108 ----- 108
Db 1441 GTTGTGTGAGACTGAGGGGGCCAGAGGTCAACACAGCTGGGCTTACGCCAAGCTTTGCT 1500
Qy 108 ----- 108
Db 1501 CTTAGCTTCGACGTGTCTCTGCTCTCCAGCGCTGTGACGTGTGCTCTGACAGGTTGC 1560
Qy 108 ----- 108
Db 1561 ACACCCCTTCGAGGCTGAAAGGCTTTGTCAAAACATTGATCTGACGAGACGCCGGTG 1620
Qy 108 ----- 108
Db 1621 CCTTGACGTGTGGGCCAGAGGCAAGTGTGAAAAGCTCCATACACAAAGAAC 1680
Qy 108 ----- 108
Db 1681 TGCAGTGAAGACGAAAGGTGCTGCTGTGATTAATAGAGCAATTTTGGACCAAGTAC 1740
Qy 109 ----- 113
Db 1741 AGTGTGTTTGAACAGGCTTTGCCCTGTCTGTTGCCCTTAGGCTTTCATGTGGTG 1800
Qy 114 GluArgSerPheLeuGluSerIleAspAspAlaLeuAlaGluLeuAlaSerLeuGlnSer 133
|||||

Db 1801 GAGAGAGCTTCTCGAGAGTCCATTGAAGAGCCTTGCTGAGAAAGCAAGCCTCAGTGC 1860
Qy 134 GlnLeuProGlu----- 137
Db 1861 CAATTGGCCAGA-GGTATTTTCCCGAGCCGACACCGAGGGAGTCAAGTCCAGGCCAGCT 1919
Qy 137 ----- 137
Db 1920 TTGCAAGAGCATGAGCTCATCTTACTTTCTTGACATTACTGGGCCAGAGCACAGCGCTT 1979
Qy 137 ----- 137
Db 1980 AGGAGCAGTTTCCGATGGGTGACACTGTTGTGGCCACAGGTGAGGGACCTCGCTGCT 2039
Qy 137 ----- 137
Db 2040 CTGCTGTGGCAGCTGCATGCTTAAGCTCAGCAGTGTGGGCGATGTGACATGAGGA 2099
Qy 137 ----- 137
Db 2100 AGGCTTTTCCAGTCTTGTGCTGTAATGGGTCCCGCCCTTGTCAACCACTGCTCAC 2159
Qy 137 ----- 137
Db 2160 ATGCGCTACTGCCACCTGACGCGCTTACCTGTGCTGTGATGGTGCATCTCCAGCTT 2219
Qy 137 ----- 137
Db 2220 TCTGGAAGAACTTAGCCTGAGCTGTAGCGAAGCTTCATCTTGCGTGTGCCATGGACC 2279
Qy 137 ----- 137
Db 2280 ATCTGTCTGTCTGCCCCGGGCCCTCGGATCTTTGCTTTAAGACAAAGGAGGCCGG 2339
Qy 137 ----- 137
Db 2340 GCATGTGGCTCAGCGCTGTAAATCCAGACACTTGGAGAGCGAGGCGGCGGATCACAA 2399
Qy 137 ----- 137
Db 2400 GGTGAGAGATCAAGACCATCTGTGCTAACATGTGTAACCCGCTCTTAAATAATAC 2459
Qy 137 ----- 137
Db 2460 AAAAAATTAGCTGGGCGTGTGTGGCGCTTGTAGTCCAGCTACTCTCGAGGCTGAA 2519
Qy 137 ----- 137
Db 2520 GCAGAGATGTGTGAACCTGGGAGGCGGAGCTTGTAGAGCCGAGATCATGCCACTG 2579
Qy 137 ----- 137
Db 2580 CACTGTAGCCTGGGGGACACAGGAGACTCATCTCAAAAAAAAAAAGATAAGG 2639
Qy 137 ----- 137
Db 2640 AGTGAAGGGCTAGGAAGTGTCCAGAGTCTGTCTGCTGCCTGACCTCTGTATATGG 2699
Qy 138 -----GlyValProGlnHisGlnLeuProGlnTyrGlnIleLeuGluArgLeu 155
|||||
Db 2700 TTGTAGGAGAGTCCCTCAGCACAGCTGCTCTCAGATACAGAAAGATCCTTGAGAGACTC 2759
Qy 156 LysThrLeuGluArgGluIleSerGlyAlaMetAlaValAlaValLeuLeuAsn 175
2760 AAGACGTTAAGAGAGGAAATTTGCGGAGGGGCGCATGCGCTGTGTGCGGCTTCTCAAC 2819
Qy 176 AsnIleLeuIleValAlaAsnVal----- 183
Db 2820 AACCAAGCTTACGTGCGCAATGTGTGAGCCCTCTGTCTCCAGGACGAGGAGACTG 2879
Qy 183 ----- 183
Db 2880 GGGAGAGGTACGCCACAGGGGTGGTGCAATTATTGACAATCTGCTTCCAGACACTTCA 2939

QY 183 ----- 183
Db 2940 CGCATTAAACCCAGGGTCTCCTGAGACGGTTGGGTATGTCCCTCTCCACAGTGAGCC 2999
QY 183 ----- 183
Db 3000 TCAGTCCCAACTGGAAGGAGAAAGACGGGATGGGAGACAGGTGCTCGCCCTTTAGTC 3059
QY 183 ----- 183
Db 3060 CCTATTCTGCTTCTTAACCTCAITCTTGACGAGTCGTATCCCATTCCTGGGTGGCTTGGGG 3119
QY 183 ----- 183
Db 3120 GCCGCTCTTCTGAAGAGGTTACCTGGCCATGAATAACCAACCACTATCCCCATGAACA 3179
QY 183 ----- 183
Db 3180 GCTCGAGATGGGCTAGGTGACAGGGACATTTGGGGTTTTGTGAGAAGACCAGACAGGTGC 3239
QY 183 ----- 183
Db 3240 AGGTTTCAGTAGAAGGACTCTGTAGAGACCCTTCTGATGATGCTGCCTTTTTTAAATAC 3299
QY 183 ----- 183
Db 3300 TCCAGTGGAAATTTTCCACAATATAGAACATAGAGTACTGATATTTAATCAACATTC 3359
QY 183 ----- 183
Db 3360 ATGGCCCATCCCAATCCAGCAGTTATCAACTGTGGCCAGCCTTCTTTGTCCCCACGT 3419
QY 183 ----- 183
Db 3420 TCTCTATTTGAAGCAATGCCAGACACTGTATCATGTTATCTGTAATATTTCACTCTG 3479
QY 183 ----- 183
Db 3480 TATCTCTAGAGAAGAGTCTGTTTTATTAGATCATAATCCTATGATTACACTGAAGAAAG 3539
QY 183 ----- 183
Db 3540 TTAAACCCAAATCTTTCGTATCACCAGCTGTCCCTAACTGGTGACTCACCCCGTTGGTCT 3599
QY 183 ----- 183
Db 3600 GAGCCTGTTTGGCCATTTCAAGTATTTCCATGTGTGAATGCTGCTTTTCCCTCTCT 3659
QY 183 ----- 183
Db 3660 GCCTTCCGGTATGCCCTATTTCTCTGTGTGTAGTCTTTGCTTAGCTGTTACACTTCT 3719
QY 184 -----GlyThrAsnArgAlaLeuLeuCysLysSerThrValAspGlyLeuGlnValThr 201
Db 3720 GCCACAGGTACAAACCGTGACITTTATGCAATCGACAGTGTGATGGTTGCAGGTGACA 3779
QY 202 GlnLeuAsnValAspHisThrThrGluAsnGluAspGluLeuPheArgLeuSerGlnLeu 221
Db 3780 CAGCTGAACGTGGACCAACACACAGAGAACCGAGATGAGCTCTTCCGTCTTTTCGCAGCTG 3839
QY 221 ----- 221
Db 3840 GGTGAGTGGGAGGTGGGAGCGGAAGCTGATCCCCATGGGCTCACCTTCGCTCGCTCTT 3899
QY 221 ----- 221
Db 3900 TGGTGTGGGTAGAGAGCGGTGTGGTAGAGGGCTGTGATCTTGGGCTCCCCAGCCAGC 3959
QY 221 ----- 221
Db 3960 CTGCTGGGTTTCATTCCCAGCACTGCCGCTTACTGGTTGGTGATTGGAACTTTAAGTTTT 4019

QY 221 ----- 221
Db 4020 CTGTGCTTCAGATCCCGCTGTGTAAATGAGGATAATAGTAACCTACCTCGGGATGTAG 4079
QY 221 ----- 221
Db 4080 GAGGATTCAGTTAGTGTGATGTGAAATGCTTCTGAGTGCCTGGCACACAGACACTT 4139
QY 221 ----- 221
Db 4140 ACTGTTATGTGTGAGACTGAAAGAGGCCAAAGATCCAGGAGGCCAGCTGCTGTCTGA 4199
QY 221 ----- 221
Db 4200 GCTGCTCTTCCCTTCCAGTGAGCTCCAGGCAGATCGGGGCACACGGGATGGCAGGG 4259
QY 221 ----- 221
Db 4260 AGACGGCAGGCAAGCTGCTCCGTGCCAGGTGGTGCCTGGAGACGGGCTCTGAGTGAGGT 4319
QY 221 ----- 221
Db 4320 GTGTGCTGATGGCGTGGGACTGAGGACACCAGGACTTTGGGTGAGTCTCTCCAG 4379
QY 221 ----- 221
Db 4380 GTGCCCTGGTGTGCTTCATTTCTTATTCAGTGGGTCTTTATTGCTTCTTCCCATGAC 4439
QY 221 ----- 221
Db 4440 TGTGTCTCTGCTCCCTTCTTTTCTTCTTCTTGTGAACAAGACAGGATTGTTGCACT 4499
QY 222 -----GlyLeuAspAlaGlyLysIleLysGlnValGlyIleIleCysGlyG 237
Db 4500 GTTTCCTCCGTAGCTTGGATGCTGGAAGATCAAGCAGGTGGGATCATCTGTGGCA 4559
QY 237 nGluSerThrArgIleGlyAspTyrLysValLysTyrGlyTyrThrAspIleAspLe 257
Db 4560 GGAGAGCACCCCGCGGATCGGGATTACAAGTTTAATATGGCTACACGACATTGACCT 4619
QY 257 uLeuSer-----AlaLalysSerLysProIleIleAlaGluProGluIleHisGlyAl 275
Db 4620 TCTCAGTAGTGTGCGCAGCCAGCTGTCCCTGTGCTTGAAGAACAGAA----- 4668
QY 275 aGlnProLeuAspGlyValThrGlyPheLeuValLeuMetSerGluGly-LeuTyrLysA 295
Db 4669 -GGTCTAGGAGGCCAAGATGGGAGGATTGATTGTGCTG-----GGAGTCGAGGCTG 4721
QY 295 laLeuGluAlaAlaHisGly-----ProGlyGlnAlaAsnGlnGluIleAla- 310
Db 4722 CAGTGAGCCGTGATCATGCCACTGCACCTCCAGCCTGGGCAACAGCAGACCCCTGTCTC 4781
QY 311 -----AlaMetIleAspThrGluPheAlaLysGlnThrSerLeuAspAlaValAlaG 328
Db 4782 AAAAAAAGTCCAGAAAGCTGGGCGACAGC---AGACTGTAGCCCCAGTGA 4838
QY 328 lnAlaValValAspArgValLysArgIleHisSerAspThrPheAlaSerGlyGlyGluA 348
Db 4839 CACATGTGCGATCAAGTAGTTCAGAGCCCGAGCAAACTCTTTCTTCTC---CCAGAGA 4895
QY 348 rg-----AlaArgPheCysProArgHisGluAspMetThrLeuLeuValArgAsn 365
Db 4896 GGTGTTTGGGCGAGTTAGGCCCGCAGCCAGCAGGAGGTGGGTACTGTAGAAACAAGGG 4955
QY 365 heGlyTyrProLeuGlyGluMet-----SerGlnProThrPro-----S 378
Db 4956 GTGGAGTTGAGTGGGCCAAAGCTGAGGGCAGCAACTAGTCCCAAGAGGAGGAGCGCA 5015
QY 378 erProAlaProAlaAlaGlyGlyArgValTyrProValSer-----ValProTyrSer 396
Db 5016 AGCCAGAGTTGGATGGGATGGACAGCAGCCAGCCTGTGGTGGTGTGCCCATGGCTGC 5075
QY 396 erAlaGlnSerThrSerLysThrSerValThrLeuSerLeuValMetProSerGln---- 414

Db	5076	AGCAGAGC-----TGTGATGATGATGATTTCCCTCTGAGGACA	5117
Qy	415	-----GlyInMeValIaNgIyAlaHisSera	424
Db	5118	CCTCCTCGACATCACCACCAGCGCTCATGGACGAATGGACAGTCTCATCTACACGG	5177
Qy	424	1aSerThrLeuAlaBgIuAlaThrProThrLeuThrAsnGlnSerProThrLeuThrLeuG	444
Db	5178	CCAGGCCCGTCAGAGCTTCGGATGTACTCTTTATGCAATTCCTACTCAGGCCCT	5237
Qy	444	InsertThrAsnThi's-ThrGlnSerSerSerSeraspIyGlyLeuPheArg	463
Db	5238	GAGCAGTCATCTGCCATCACCACCCAGAGGCATTTTCATAGAGGAACCTAGGCGATGGAGA	5297
Qy	464	SerArgProAlaHisSerLeuProProGly	473
Db	5298	CATCAGCAGCCAC-----CCTGCT	5318
RESULT 9			
AL36984			
ID	AL36984	standard; DNA; 16877 BP.	
XX	AC	AA36984;	
XX	DT	08-JAN-2002 (first entry)	
DE	XX	Human musculoskeletal system related polynucleotide SEQ ID NO 3349.	
XX	KW	Cytostatic; immunosuppressive; nocotropic; neuroprotective; antiviral;	
KW	KW	antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;	
KW	KW	vulnerable; anticonvulsant; antibacterial; antifungal; antiparasitic;	
KW	KW	cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;	
XX	XX	neurological disease; infection; human; secreted protein;	
XX	XX	musculoskeletal system; ds.	
OS	XX	Homo sapiens.	
XX	PN	WO20015367-A1.	
XX	PD	02-AUG-2001.	
XX	PF	17-JAN-2001; 2001WO-US01338.	
XX	PR	31-JAN-2000; 2000US-0179065.	
PR	PR	04-FEB-2000; 2000US-0180628.	
PR	PR	24-FEB-2000; 2000US-0184664.	
PR	PR	02-MAR-2000; 2000US-0186350.	
PR	PR	16-MAR-2000; 2000US-0189874.	
PR	PR	17-MAR-2000; 2000US-0190076.	
PR	PR	18-APR-2000; 2000US-0198123.	
PR	PR	19-MAY-2000; 2000US-0205515.	
PR	PR	07-JUN-2000; 2000US-0209467.	
PR	PR	28-JUN-2000; 2000US-0214886.	
PR	PR	30-JUN-2000; 2000US-0215135.	
PR	PR	07-JUL-2000; 2000US-0216647.	
PR	PR	07-JUL-2000; 2000US-0216880.	
PR	PR	11-JUL-2000; 2000US-0217487.	
PR	PR	11-JUL-2000; 2000US-0217496.	
PR	PR	14-JUL-2000; 2000US-0218230.	
PR	PR	26-JUL-2000; 2000US-0220963.	
PR	PR	26-JUL-2000; 2000US-0220964.	
PR	PR	14-AUG-2000; 2000US-0224518.	
PR	PR	14-AUG-2000; 2000US-0224519.	
PR	PR	14-AUG-2000; 2000US-0225213.	
PR	PR	14-AUG-2000; 2000US-0225214.	
PR	PR	14-AUG-2000; 2000US-0225266.	
PR	PR	14-AUG-2000; 2000US-0225267.	
PR	PR	14-AUG-2000; 2000US-0225268.	
PR	PR	14-AUG-2000; 2000US-0225270.	
PR	PR	14-AUG-2000; 2000US-0225447.	
PR	PR	14-AUG-2000; 2000US-0225757.	
PR	PR	14-AUG-2000; 2000US-0225758.	

[illegible]

PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251038.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
PA (HUMA-) HUMAN GENOME SCI INC.
XX

Rosen CA, Barash SC, Ruben SM;
WPI; 2001-451937/48.

Isolated polypeptide for treating, preventing and/ or prognosing
disorders related to the musculoskeletal system including
musculoskeletal cancers and also for testing and detection e.g.
diagnosis -

Example 2; SEQ ID NO 3349; 781pp + Sequence Listing; English.

The invention relates to novel genes (AAL34669-AAL37666) and proteins (ABB03087-ABB04109) associated with the musculoskeletal system useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemia; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from wipo at ftp.wipo.int/pub/published_pct_sequences.

Sequence 16877 BP; 3563 A; 4508 C; 4516 G; 4290 T; 0 other;

Alignment Scores:

Pred. No.: 2,27e-30 Length: 16877
Score: 543.50 Matches: 295

Percent Similarity: 18.86% Conservative: 43
Best Local Similarity: 16.46% Mismatches: 107
Query Match: 21.07% Indels: 1348
DB: 22 Gaps: 19

US-09-830-144-4 (1-504) x AAL36984 (1-16877)

QY 12 GluGlnGlnProSerTrrThrAspAspLeuProLeuCysHisLeuSerGlyValGlySer 31
DB 1 GAGCAGCAGCCNAGCTGGACAGATGACTGCTCTCTCCACCTCTCTGGGGTTGGCTCA 60
QY 32 AlaSerAsnArgSerTyrSerAlaAspGlyLysGlyThrGluSerHisProGluAsp 51
DB 61 GCCTCCAAACCGCAGCTACTCTGCTGATGGCAAGGSCACTGAGAGCCACCCCGCAGAGAC 120
QY 52 SerTrpLeuLysPhe----- 56
DB 121 AGCTGGCTCAAGTTTCAGGTGTGTGTGCCAGCATTTCTGTGTGGCCCGGGAGTTGGTT 180
QY 56 ----- 56
DB 181 GGTTTGCAAGCAAGAAAGACACCGACCTTGCAGCTTTCTCGTATGGGCTTGCAGTGAC 240
QY 56 ----- 56
DB 241 ATGTGGCCCGTGAGAGGTGGCTCTGCTGTGTCTTGCCAAAGGCCCTGCTCTGATGGGT 300
QY 56 ----- 56
DB 301 AGCGTGAGCATGGGAGGAGGGGTATCCCAAGATGTATAGCCAGAGTGAATGATGGCTA 360
QY 56 ----- 56
DB 361 AAGCAGGGGGACCCAGGAGGGCCCTGAAGCTGCAGCTGTGTGCTTTAGTCTCCCCCA 420
QY 56 ----- 56
DB 421 ATTCCTTTCCCTCCCTCCCTCGTGAGACCCCTGGTCTCAGGCCCTCCCTCTGCC 480
QY 57 -----ArgSerGluAsnAsnCysPheLeuTyrGlyValPheAsnGlyTyrAsp 72
DB 481 TCTCCCTCTCCAGGAGGTGAGAACAACTGCTTCTGTATGGGGTCTTCAACGGCTATGAT 540
QY 73 GlyAsnArgValThrAsnPheValAlaGlnArgLeuSerAlaGluLeuLeuGlyGln 92
DB 541 GGCACCGAGTGACCACTTCTGTGCCCGCAGCGCTGTCCGACAGAGCTCTGCTGGCCAG 600
QY 93 LeuAsnAlaGluHisAlaGluAlaAspValArgArgValLeuLeuGln----- 108
DB 601 CTGAATGCCGAGCAGCGCGAGCGGCGATGTGCGGCGTGTGCTGTGCAGGTAAATGGTCCG 660
QY 108 ----- 108
DB 661 GGGCCAAAGTGAACCCAGCCACATCATGTCCCCACCCCAAGGCTTGGGCCCTGCACCTC 720
QY 108 ----- 108
DB 721 TAGCATGTTGCCAGGGTGTGTGAAGATCTGCGCGGCCCTTCAACCCAGTAGAGGAGC 780
QY 108 ----- 108
DB 781 AGCTCCAGCGTAGGCCCCCCCCCAACAGGAGTCCAGGACCCAGCCAGCAGCAGCAGG 840
QY 108 ----- 108
DB 841 CCTTCAATGGCTGGCAGAGGAGGTACTGCTGGAGACGGGGGATTATTAGGATGGAGCTT 900
QY 108 ----- 108
DB 901 GGAGAGAGGTGTGAGGTGGGAGCAGGGCAAGCCCTGTAGAAATGGGGTCATTTAGAGCT 960
QY 108 ----- 108

Db 961 ACCCCCTTCTTCTATGTGTGTCAGTGTGCTCCAGTGCAGAGGACCATCCCTGG 1020
QY 108 ----- 108
Db 1021 CGTCCAGAAAGACCTTGCCCTCCCTTTCTGAGGGGCGCGCCCTCATTTGACTGTTCC 1080
QY 108 ----- 108
Db 1081 ACCTTCTGCTTGCGCTGTGGCAGTTTTCCTCTGGCATCTGCTTTGAGAGCATGTCT 1140
QY 108 ----- 108
Db 1141 CAGGCCCATTTGATGAGAAGATGGGCTTCTGTTCCCGAGAGGGGTGTGCCAGCTT 1200
QY 108 ----- 108
Db 1201 TTCTGCTCTTACGACCTCAGGCTTCATTGGCAGTATTTCTCAGCAGATCTCACACAG 1260
QY 108 ----- 108
Db 1261 GGGAGAGGTGTCAGTGTCAACAGTGTCTGTGGGCTGTGGGTTTGACAGAAAGCTTCCC 1320
QY 108 ----- 108
Db 1321 AGAAGAAGTTGTATATCATACAGCTGTCAATAGGCTGTGGCAGTTTCTGACAGTT 1380
QY 108 ----- 108
Db 1381 GCCTTGCTGAGTGCAGGTTTGGACTGTCTTACTTAACCTTGAAACACTTAACTTCC 1440
QY 108 ----- 108
Db 1441 GTTGTGTGAGACTGAGGGGGCCAGAGTCAACACAGCTGGGCTTACGCCAAGCTTTGCT 1500
QY 108 ----- 108
Db 1501 CTTAGCTTCGCACTGTCTCTGCTCCAGCCGTCGTGAGTGTGCTCTGCAAGGTGC 1560
QY 108 ----- 108
Db 1561 ACACCCCTTCGAGGCTGAAGGGCTTTGTCAAGACATTGATCTGCAGAAAGCCGGTG 1620
QY 108 ----- 108
Db 1621 CTTGAGTGTGTCGGCCAGAGGAGCACTGACATGTGAAGCTCCATCACAAAGAAC 1680
QY 108 ----- 108
Db 1681 TGCAGTGAMGACAGAAAGCTGCTGCTGTGATTATAGAGACATTTTGGCACAGTGAC 1740
QY 109 ----- 113
Db 1741 AGTGTGTTTTGAACCAAGCTTTGCCCTGTCTGTGTCCCTCAGGCTTTCATGTGGTG 1800
QY 114 GluArgSerPheLeuGlnSerIleAspAspAlaLeuAlaGlnIleAlaSerLeuGlnSer 133
Db 1801 GAGAGGAGCTTCTCGGATCCATTGACGAGCCTTGTGCTGAGAAAGGCAAGCCTCCAGTGG 1860
QY 134 GlnLeuProGlu----- 137
Db 1861 CAATTGGCAGA-GGTAAATTTCCCGAGCCGACCCAGGGAGTCAAGTCCAGGCCCAAGCT 1919
QY 137 ----- 137
Db 1920 TTGCAAGAGACATGATCTACTTCTTGACATTACTGGGCCAGAGCAAGGCGTT 1979
QY 137 ----- 137
Db 1980 AGGAGCAGTTCTGATGGGTGACACTGTGTGTGGCCACAGGTGAGGACCTCGCTGCT 2039
QY 137 ----- 137
Db 2040 CTGCTGTGGCAGCTGCATGCTAAAGCTCAGCAGTGTCTGGGAGTGGTGACATGAGA 2099

QY 137 ----- 137
Db 2100 AGGCTTTTCCAGTCTCTTGTCTTGATGGGTCCCGGCCCTTGTCAACCACTGCTCAC 2159
QY 137 ----- 137
Db 2160 ATGGCTACTGCACCTGCAGGCTTACCTGTGTGTGTGAGTGTGGTATCTCCAGCTT 2219
QY 137 ----- 137
Db 2220 TCTGAAGAACTTAGCTGAGCTGTAGCGAAGCTTCATCTTGCTGTGCAATGACC 2279
QY 137 ----- 137
Db 2280 ATCTGTGTGTGTCCCCGGGCCCTTGGATCTTGTCTTAAAGCAAAAGGAGGCCGG 2339
QY 137 ----- 137
Db 2340 GCATGTGTGCTCAGGCTGTATATCCAGACACTTGGGAGGCGGAGGCGGATCACA 2399
QY 137 ----- 137
Db 2400 GGTGAGATCAAGACCATCTGGCTAACATGTGTAACCCCGTCTTAAATAAC 2459
QY 137 ----- 137
Db 2460 AAAAAATTAGCTGGGCGTGTGTGGGCGCTTGTGTAGTCCAGCTACTCTCGAGGCTGAA 2519
QY 137 ----- 137
Db 2520 GCAGAGATGTGTGTGAACCTGGGAGGCGGAGCTTGTAGTGAACGAGATCATGCCACTG 2579
QY 137 ----- 137
Db 2580 CACTGTAGCTGGGGGACACAGCGAGACTCATCTCAAAAAAAAAAAGTAAAGG 2639
QY 137 ----- 137
Db 2640 AGTGAAGGCGTAGAAGATGCTCCAGAGTTCTGTCTGACCCTCTGTTATGG 2699
QY 138 ----- 155
Db 2700 TTGTAGGAGTCCCTCAGCACAGCTGCTCTCAGTATAGAAAGATCCTTGAGAGCTC 2759
QY 156 LysThrLeuGlnArgGlnIleSerGlyAlaMetAlaValAlaValLeuLeuAsn 175
Db 2760 AAGCGTTAGAGAGGAAATTTGCGGAGGGGCATGGCCGTGTGGGGTCTTCTCAAC 2819
QY 176 AsnLysLeuTyrValAlaAsnVal----- 183
Db 2820 AACAACTTAACGTTCGCCCAATGTGTGAGCCCTCTGTGCCAGGGGAGGAGACTG 2879
QY 183 ----- 183
Db 2880 GGGAGAGTCAAGCCACAGGGGTGGTGCATTATTGCAATCTGCTTTCCAGACACTTCA 2939
QY 183 ----- 183
Db 2940 CGCACTTAAACCAAGGCTCTCTGAGACCGTTGGTATGCTCCCTCCACAGTGAGCC 2999
QY 183 ----- 183
Db 3000 TCAGTCCCACTGGAAGGAGAAAGACGGGATGGAGACAGGTGTCCTGGCTTTAGTC 3059
QY 183 ----- 183
Db 3060 CCTATTCTGCTTTTAACTATTCTGACAGAGTGTATCCCATTTCTGGGTGGCTTGGGG 3119
QY 183 ----- 183
Db 3120 GCCGCTGCTTGTGAAGAGTTTACCTGGCCATGAATAACCAACCAAGTATCCCATGAACA 3179

QY 183 ----- 183
Db 3180 GGTGAGATGGGGCTAGGTGACAGGACATTTGGGGTTTGTGAGAGACAGAGGTGC 3239
QY 183 ----- 183
Db 3240 AGGTTTCAGTAGAAGGACTCTGTAGAGACCTTCTGATGATGCTGCTTTTAAATAC 3299
QY 183 ----- 183
Db 3300 TCCAGTGGAAAAATTTCCACAATATAGAACAAATAGAGTGAATGATATTAATGAACATTC 3359
QY 183 ----- 183
Db 3360 ATGCCCCATCCCAATCCAGCAGTTATCAACTGTGGCAGCCCTCTTTGTCCCCACGT 3419
QY 183 ----- 183
Db 3420 TCTCTATTTGAAGCAATGCCAGACACTGTATCATGTTATCTGTAATAATTTTCAGTCTG 3479
QY 183 ----- 183
Db 3480 TATCTCTAGAGAAGGCTCTGTTTATTAAGATCATATCCTATGATTAACACTGAAAAAG 3539
QY 183 ----- 183
Db 3540 TTAAACCCCAATCTTTCGTATCACCAGCTGTCCCTAACTGGTGACTCACCCCGTTGGTCT 3599
QY 183 ----- 183
Db 3600 GAGCCTGTTTGGCCATTTCCAGTATTTCCATGTGTGAAATGCTGCTTTTCCCTCTCT 3659
QY 183 ----- 183
Db 3660 GCCTCCCGGTATGCCCTATTCTCTGTGTGTAGTCTTTGTGTAGCTGTTCACATTCT 3719
QY 184 ----- GlyThrAsnArgAlaLeuLeuCysLysSerThrValaspGlyLeuGlnValThr 201
Db 3720 GCCACAGGTACAAACCGTGCACTTTATGCAAAATCGACAGTGGATGGTTGCAGGTGACA 3779
QY 202 GlnLeuAsnValAspHisThrThrGlnuGlnuAspGluLeuPheArgLeuSerGlnLeu 221
Db 3780 CAGCTGAACGTGGACACACACAGAGAACAGGATGAGCTCTTCCTGCTTTCGCACGTG 3839
QY 221 ----- 221
Db 3840 GGTGAGTGGGAGAGTGGGAGCGGAAGCTGATCCCATGGGCTCACCTTCGCGTGCCTT 3899
QY 221 ----- 221
Db 3900 TGGTGTGGGTAGAGAGCGGTGTGTAGAGGGCTGTGATCTTGGGCTCCCCAGCCAGC 3959
QY 221 ----- 221
Db 3960 CTGCTGGGGTTTCATCCAGCACTGCCGCTTACTGGTTGGTGAATTGGAACCTTAAGTTTT 4019
QY 221 ----- 221
Db 4020 CTGTGCTCAGATCCCCCGCTGTGTAATGAGGATAATAGTAACCTACCTCGGGATGCTAG 4079
QY 221 ----- 221
Db 4080 GAGGATTCAGTTAGTGCATGTGAAATGCTTCTGGAGTGCCTGGGCACAGAGACACTT 4139
QY 221 ----- 221
Db 4140 ACTGTTATTGGTGGAGACTGAAAGAGGCCAAAGAGTCCAGGAGCCAGCTGTGCTGA 4199
QY 221 ----- 221
Db 4200 GCTGCTCCTTCCCTTCCAGTGAGCTCCAGGCAGATGTGGGGCACACGGGATGGCAGGG 4259
QY 221 ----- 221

Db 4260 AGACGGCAGCAAGCTGCTCCGTGCCAGGTGGTGCTCTGGAGGACGGGCTCTGAGTGAGGT 4319
QY 221 ----- 221
Db 4320 GTGTCGTGATGGGCGTGGGACTGAGGACACACAGGAGACTTTGGGTCAGCTGCTCTCCAG 4379
QY 221 ----- 221
Db 4380 GTGCCCTGGTGTGTCTTCTTCTTCTTATTCAGTGGGTCTTATTGCTCTTCTTCCCATGAC 4439
QY 221 ----- 221
Db 4440 TGTGTCTCTGTCCCTTCTTTTGTCTCTTGTGAAACAAGACAGGATGTTTGCACCT 4499
QY 222 ----- GlyLeuAspAlaGlyLysIleLysGlnValGlyIleLeuCysGlyG 237
Db 4500 GTTTCCTCCGTAGGCTTGGATGCTGGAAGATCAAGCAGTGGGATCATCTGTGGGCA 4559
QY 237 nGluSerThrArgArgIleGlyAspTyrLysValLysTyrGlyTyrThrAspIleAspLe 257
Db 4560 GGAGAGCACCCGCGGATCGGGATTACAAGGTTAAATATGGCTACACGGACATTGACCT 4619
QY 257 uLeuSer-----AlaAlaLysSerLysProIleIleAlaGluProGluIleHisGlyAl 275
Db 4620 TCTCAGGTAGGTGCGCAGCCAGCTGTCCCTGTGCTTGAAAGAACAGAA----- 4668
QY 275 aGlnProLeuAspGlyValThrGlyPheLeuValLeuMetSerGluGly-LeuTyrLysA 295
Db 4669 -GGTCTAGGAGGCCAAGATGGGAGGATTCATGTTGCTG-----GGAGTCGAGGCTG 4721
QY 295 laLeuGluAlaAlaHisGly-----ProGlyGlnAlaAsnGlnGluIleAla- 310
Db 4722 CAGTGAGCGGTGATGCCACTGCACTCCAGCTGGGCAACAGAGCAGACCTGTCTC 4781
QY 311 -----AlaMetIleAspThrGluPheAlaLysGlnThrSerLeuAspAlaValAlaG 328
Db 4782 AAAAAAAGGTCCAGAAAGCTGGGCGCAGACGC---AGACTGTAGGCCCCAGTGA 4838
QY 328 lnAlaValValAspArgValLysArgIleHisSerAspThrPheAlaSerGlyGlyGlu 348
Db 4839 CACATGTCAGTCAAGTAGTTCAGAGCCAGCAAACTCTTCTTCCTC---CCAGAGA 4895
QY 348 rg-----AlaArgPheCysProArgHisGluAspMetThrLeuLeuValArgAsnP 365
Db 4896 GTTGGTTTGGGCGAGTTAGGCCCAGCAGCAGGAGGTGGGTACTGTAGAAAAACAAGG 4955
QY 365 heGlyTyrProLeuGlyGluMet-----SerGlnProThrPro-----S 378
Db 4956 GTGGAGTTGAGCTGGGCCAAAGCTGAGGCGAGCAACTAGTCCCAAGGAGGAGCGCGCA 5015
QY 378 erProAlaProAlaAlaGlyArgValTyrProValSer-----ValProTyrSerS 396
Db 5016 AGCCAGAGTTGGATGGATGGACAGACAGGCGCTGTGTGCGTGTGCGCTGCCATGGCTGC 5075
QY 396 erAlaGlnSerThrSerLysThrSerValThrLeuSerLeuValMetProSerGln---- 414
Db 5076 AGCAGAGG-----TGTGAGATGAATTGTATTCTCCCTCTGAGGACA 5117
QY 415 -----GlyGlnMetValAsnGlyAlaHisSerA 424
Db 5118 CCTCCTCTGGACATCACCCAGCGCTCATGGCAGCAGAAATGGGCGAGTCATCATCACGG 5177
QY 424 laSerThrLeuAspGluAlaThrProThrLeuThrAsnGlnSerProThrLeuThrLeuG 444
Db 5178 CCAAGCCGCTGAGAGCTCTGGATGTACCTTCTTTTATGATTCCTCTACTCAGGCCCTCT 5237
QY 444 lnSerThrAsnThrHis-ThrGlnSerSerSerSerSerSerSerSerGlyGlyLeuPheArg 463
Db 5238 GAGGCACTCATCTGCGATCACCCAGAGGCAATTTTCAAGAGGAAACTGAGGATGGGAGA 5297
QY 464 SerArgProAlaHisSerLeuProProGly 473
QY 473 -----

Db 5298 CATCAGCAGCCAC-----CTGGT 5318
RESULT 10
AAFI5895
ID AAFI5895 standard; cDNA; 696 BP.
XX
AC AAFI5895;
XX
DT 13-MAR-2001 (first entry)
XX
DE Human prostate cancer antigen nucleotide sequence SEQ ID NO:330.
XX
KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KW neuroprotective; cytosolic; cardioactive; immunomodulatory; muscular;
KW vulnerrary; gastrointestinal; nephrotoxic; antiinfective; gynaecological;
KW antibacterial; gene therapy; neural; immune; reproductive; renal;
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
KW wound; infectious disease; ss.
XX
OS Homo sapiens.
XX
PN WO200055174-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US05988.
XX
PR 12-MAR-1999; 99US-0124270.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.
XX
PI Rosen CA, Ruben SM;
XX
XX WPI, 2000-587513/55.
DR P-PSDB; AAB56692.
XX
XX Prostate cancer associated gene sequences, referred to as prostate
PT cancer antigens, useful for treatment, prevention, and diagnosis of
XX disorders such as prostate cancer -
XX
XX Claim 1; Page 837; 2338pp; English.
XX
XX AAFI5566 to AAFI6505 encode the human prostate cancer associated
CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC The prostate cancer antigens can have neuroprotective, cytosolic,
CC cardioactive, immunomodulatory, muscular, vulnerrary, gastrointestinal,
CC nephrotoxic, antiinfective, gynaecological and antibacterial activities,
CC and can be used in gene therapy. The prostate cancer antigen
CC polynucleotides may be used for detection of prostate cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The prostate cancer antigens may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AAFI6506 to AAFI6514 to
CC AAB57303 represent sequences used in the exemplification of the present
CC invention.
XX
SQ Sequence 696 BP; 143 A; 237 C; 187 G; 120 T; 9 other;
XX
Alignment Scores:
Pred. No.: 5.27e-22 Length: 696
Score: 409.00 Matches: 78
Percent Similarity: 98.75% Conservative: 1
Best Local Similarity: 97.50% Mismatches: 1
Query Match: 15.85% Indels: 0
DB: 21 Gaps: 0
US-09-830-144-4 (1-504) x AAFI5895 (1-696)
QY 425 SerThrLeuSpGluAlaThrProThLeuThrAsnGlnSerProThrLeuThrLeuGln 444
DB 13 GCCACCTGAGACGACCCCACTCACCACCAAGCCCGACCTTAACCTTGAG 72

QY 445 SerThrAsnThiHisThrGlnSerSerSerSerSerAspGlyGlyLeuPheArgSer 464
DB 73 TCACCCACACGACACGACGAGCGAGCTTCAGCTTACGAGGAGGCTCTTCGCTTC 132
QY 465 ArgProAlaHisSerLeuProProGlyGluAspGlyArgValGluProTyValAspPhe 484
DB 133 CGGCCCGCCCATCGCTCCGCGCTGCGAGACGAGTCTGTTGAGCCCTTAGTGGAATTT 192
QY 485 AlaGluPheTyArgLeuTrpSerValAspHisGlyGlnGlnSerValValThrAlaPro 504
DB 193 GCTGAGTTTACCGCCTTGAGCGTGGACCATGGCGAGAGCGTGTGACACACCG 252
RESULT 11
AB447246
ID AB447246 standard; DNA; 211 BP.
XX
AC AB447246;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human breast cell single exon nucleic acid probe #5941.
XX
KW Human; microarray; single exon probe; gene expression; breast;
KW disease; cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157271-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JUN-2001; 2001WO-US00662.
XX
PR 04-FEB-2000; 2000US-0180312.
XX
PR 26-MAY-2000; 2000US-0207456.
XX
PR 30-JUN-2000; 2000US-0608408.
XX
PR 03-AUG-2000; 2000US-0632366.
XX
PR 21-SEP-2000; 2000US-0234687.
XX
PR 27-SEP-2000; 2000US-0236359.
XX
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-496933/54.
XX
XX New spatially-addressable set of single exon nucleic acid probes,
PT useful for measuring gene expression in sample derived from human
PT breast, comprises number of single exon nucleic acid probes -
XX
XX Claim 4; SEQ ID NO 5941; 327bp + sequence listing; English.
XX
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting
CC the probes with a collection of detectably labelled nucleic acids
CC derived from mRNA of human breast, and then measuring the label
CC bound to each probe of the microarray. The probes are useful for
CC verifying the expression of regions of genomic DNA predicted to
CC encode proteins. They are useful for gene discovery, and for
CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater
CC diversity of probes for measuring gene expression, with far less bias
CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence. The
CC present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 211 BP; 43 A; 70 C; 66 G; 32 T; 0 other;
Alignment Scores:
Pred. No.: 2.38e-19 Length: 211
Score: 365.00 Matches: 70
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.15% Indels: 0
DB: 22 Gaps: 0

US-09-830-144-4 (1-504) x ABA47246 (1-211)

QY 312 MetIleAspThrGluPheAlaLysGlnThrSerLeuAspAlaValAlaValVal 331
Db 1 ATGATTGACACTGAGTTTGCCAGCAGACCTCCCTGGACGAGTGGCCCGTCGTG 60
QY 332 AspArgValLysArgIleHisSerAspThrPheAlaSerGlyGlyGluArgAlaArgPhe 351
Db 61 GACCGGTGAAGCCATCCACAGCAGACCTTCGCCAGTGGTGGGAGCGTGCAGGTTT 120
QY 352 CysProArgHisGluAspMetThrLeuValArgAsnPheGlyTyrProLeuGlyGlu 371
Db 121 TGCCCCCGCAGCAGGACATGACCTCTAGTGAGGAACCTTTGGCTACCGCTGGGCGAA 180
QY 372 MetSerGlnProThrProSerProAlaPro 381
Db 181 ATGAGCCAGCCACACCGAGCCAGCCCCA 210

RESULT 12

ABAE65131

ID ABA65131 standard; DNA; 211 BP.

XX AC ABA65131;

XX DT 01-FEB-2002 (first entry)

XX DE Human foetal liver single exon nucleic acid probe #13436.

XX KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

XX OS Homo sapiens.

XX PN WO200157277-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00669.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-483447/52.

XX PT Human genome-derived single exon nucleic acid probes useful for

XX PT analyzing gene expression in human fetal liver -

XX PS Claim 4; SEQ ID NO 13436; 639pp + sequence listing; English.

XX CC The invention relates to a single exon nucleic acid probe for

XX CC measuring human gene expression in a sample derived from human foetal

XX CC liver. The single exon nucleic acid probes may be used for predicting,

XX CC measuring and displaying gene expression in samples derived from human

XX CC foetal liver. The present sequence is a single exon nucleic acid

XX CC probe of the invention.

CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 211 BP; 43 A; 70 C; 66 G; 32 T; 0 other;
Alignment Scores:

Pred. No.: 2.38e-19 Length: 211
Score: 365.00 Matches: 70
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.15% Indels: 0
DB: 22 Gaps: 0

US-09-830-144-4 (1-504) x ABA65131 (1-211)

QY 312 MetIleAspThrGluPheAlaLysGlnThrSerLeuAspAlaValAlaValVal 331
Db 1 ATGATTGACACTGAGTTTGCCAGCAGACCTCCCTGGACGAGTGGCCCGTCGTG 60
QY 332 AspArgValLysArgIleHisSerAspThrPheAlaSerGlyGlyGluArgAlaArgPhe 351
Db 61 GACCGGTGAAGCCATCCACAGCAGACCTTCGCCAGTGGTGGGAGCGTGCAGGTTT 120
QY 352 CysProArgHisGluAspMetThrLeuValArgAsnPheGlyTyrProLeuGlyGlu 371
Db 121 TGCCCCCGCAGCAGGACATGACCTCTAGTGAGGAACCTTTGGCTACCGCTGGGCGAA 180
QY 372 MetSerGlnProThrProSerProAlaPro 381
Db 181 ATGAGCCAGCCACACCGAGCCAGCCCCA 210

RESULT 13

ABAE65131

ID ABA32233 standard; DNA; 211 BP.

XX AC ABA32233;

XX DT 23-JAN-2002 (first entry)

XX DE Probe #10699 for gene expression analysis in human heart cell sample.

XX KW Human; gene expression; heart; microarray; vascular system; probe;

XX KW cardiovascular disease; hypertension; cardiac arrhythmia;

XX KW congenital heart disease; ss.

XX OS Homo sapiens.

XX PN WO200157274-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00666.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-488899/53.

XX PT Single exon nucleic acid probes for analyzing gene expression in human

XX PT hearts -

XX PS Claim 4; SEQ ID No 10699; 530pp; English.

XX PS

CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 211 BP; 43 A; 70 C; 66 G; 32 T; 0 other;

Alignment Scores:

Pred. No.:	2,38e-19	Length:	211
Score:	365.00	Matches:	70
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	14.15%	Indels:	0
DB:	22	Gaps:	0

US-09-830-144-4 (1-504) x ABA32233 (1-211)

OY 312 MetIleAspThrGluPheAlaIysGlnThrSerLeuAspAlaValAlaGlnAlaVal 331

DB 1 ATGATTGACACTGAGTTTGCCAAAGACGACCTCCCTGAGCGAGTGCCAGGCCGTCTG 60

OY 332 AsparGValIysArgIleHisSerAspThrPheAlaSerGlyGluArgAlaArgPhe 351

DB 61 GACCGGGTGAAGCGATCCACAGCCACCTTCGCCAGTGGGGAGCCGTGCCAGTTTC 120

OY 352 CysProArgHisGluAspMetThrLeuLeuValArgAsnPheGlyTyrProLeuGlyGlu 371

DB 121 TGCCCCCGGACAGAGCATGACCTCGTAGAGGAACCTTGCTACCCGCTGGCGAA 180

OY 372 MetSerGlnProThrProSerProAlaPro 381

DB 181 ATGAGCCAGCCACACCGAGCCAGCCCA 210

RESULT 14

AAK13550
ID AAK13550 standard; DNA; 211 BP.

XX AAK13550;

DT 05-NOV-2001 (first entry)

XX Human brain expressed single exon probe SEQ ID NO: 13541.

XX Human; brain expressed exon; gene expression analysis; probe;

KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;

KW epilepsy; cancer; ss.

XX Homo sapiens.

PN WO200157275-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00667.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483446/52.

XX Single exon nucleic acid probes for analyzing gene expression in human

PT brains -

PS Example 4; SEQ ID NO: 13541; 650bp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in the human

CC brain. They can be used to measure gene expression in brain cell samples,

CC which may enable the diagnosis and improved treatment of nervous system

CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,

CC epilepsy and cancers. The present sequence is one of the probes of the

CC invention.

XX Sequence 211 BP; 43 A; 70 C; 66 G; 32 T; 0 other;

Alignment Scores:

Pred. No.:	2,38e-19	Length:	211
Score:	365.00	Matches:	70
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	14.15%	Indels:	0
DB:	22	Gaps:	0

US-09-830-144-4 (1-504) x AAK13550 (1-211)

OY 312 MetIleAspThrGluPheAlaIysGlnThrSerLeuAspAlaValAlaGlnAlaVal 331

DB 1 ATGATTGACACTGAGTTTGCCAAAGACGACCTCCCTGAGCGAGTGCCAGGCCGTCTG 60

OY 332 AsparGValIysArgIleHisSerAspThrPheAlaSerGlyGluArgAlaArgPhe 351

DB 61 GACCGGGTGAAGCGATCCACAGCCACCTTCGCCAGTGGGGAGCCGTGCCAGTTTC 120

OY 352 CysProArgHisGluAspMetThrLeuLeuValArgAsnPheGlyTyrProLeuGlyGlu 371

DB 121 TGCCCCCGGACAGAGCATGACCTCGTAGAGGAACCTTGCTACCCGCTGGCGAA 180

OY 372 MetSerGlnProThrProSerProAlaPro 381

DB 181 ATGAGCCAGCCACACCGAGCCAGCCCA 210

RESULT 15

AAK39289
ID AAK39289 standard; DNA; 211 BP.

XX AAK39289;

DT 06-NOV-2001 (first entry)

XX Human bone marrow expressed single exon probe SEQ ID NO: 13846.

XX Human; bone marrow expressed exon; gene expression analysis; probe;

KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.

XX Homo sapiens.

PN WO200157276-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00668.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

```
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX DR WPI; 2001-488900/53.
XX
XX PT Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human bone marrow -
XX
XX Example 4; SEQ ID NO: 13846; 658pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX bone marrow. They can be used to measure gene expression in bone marrow
XX samples, which may enable the improved diagnosis and treatment of cancers
XX such as lymphoma, leukaemia and myeloma. The present sequence is one of
XX the probes of the invention.
XX
XX SQ Sequence 211 BP; 43 A; 70 C; 66 G; 32 T; 0 other;

Alignment Scores:
Pred. No.: 2,38e-19 Length: 211
Score: 365.00 Matches: 70
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.15% Indels: 0
DB: 22 Gaps: 0

US-09-830-144-4 (1-504) x AAK39289 (1-211)

QY 312 MetIleAspThrGluPheAlaLysGlnThrSerLeuAspAlaValAlaGlnAlaValVal 331
Db 1 ATGATTGACACTGAGTTTGCCAAAGCAGACCTCCTGGACGACGTGCCCGCCGTCGTG 60
QY 332 AspArgValLysArgIleHisSerAspThrPheAlaSerGlyGlyGluArgAlaArgPhe 351
Db 61 GACCGGGTGAAGCGCATCCACAGCGACACCTTCGCCAGTGGTGGGAGCGTCCAGGTTTC 120
QY 352 CysProArgHisGluAspMetThrLeuValArgAsnPheGlyTyrProLeuGlyGlu 371
Db 121 TGCCCCCGGCACGAGACATGACCTGTAGTGAGGAACCTTTGGCTACCCGCTGGGCGAA 180
QY 372 MetSerGlnProThrProSerProAlaPro 381
Db 181 ATGAGCCAGCCACACCGAGCCCGCCCA 210
```

Search completed: December 9, 2002, 23:14:45
Job time : 284.44 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 9, 2002, 22:59:44 ; Search time 1573.89 Seconds
(without alignments)
5186.195 Million cell updates/sec

Title: US-09-830-144-4

Perfect score: 2580

Sequence: 1 MAARRLSLQSQPSTWDD.....AEFYRLMSVDHGEOVVTAP 504

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlp
-Q=/cgr2_1/USPTO.spool/US09830144/runat_04122002_141353_2264/app query.fasta_1.1422
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -DOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09830144 @CGN 1 1 2441 @runat_04122002_141353_2264 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_htc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: gb_gss.*
18: em_gss_hum.*
19: em_gss_inv.*
20: em_gss_pln.*
21: em_gss_vrt.*
22: em_gss_fun.*
23: em_gss_mam.*
24: em_gss_mus.*
25: em_gss_other.*
26: em_gss_pro.*
27: em_gss_rod.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1500.5	58.2	1036	13	BM560774	BM560774 AGENCOURT
2	1465	56.8	1086	14	BM906573	BM906573 AGENCOURT
3	1376.5	53.4	883	14	BQ929862	BQ929862 AGENCOURT
4	1352	52.4	863	14	BQ322049	BQ322049 AGENCOURT
5	1289	50.0	867	14	BQ431917	BQ431917 AGENCOURT
6	1280.5	49.6	847	13	BI915996	BI915996 603184416
7	1280	49.6	892	14	BQ644850	BQ644850 AGENCOURT
8	1266.5	49.1	991	14	BQ645086	BQ645086 AGENCOURT
9	1212.5	47.0	784	14	BQ442448	BQ442448 UI-M-EVO-
10	1188.5	46.1	917	14	BQ917606	BQ917606 AGENCOURT
c 11	1166	45.2	815	13	BI181306	BI181306 UNL-P-FN-
12	1165.5	45.2	910	12	BG424017	BG424017 602447479
13	1150	44.6	714	14	BQ179703	BQ179703 UI-M-EWO-
14	1135	44.0	690	10	AW247232	AW247232 2820685.5
15	1129	43.8	702	12	BG715206	BG715206 602675668
16	1106	42.9	840	13	BI918524	BI918524 603176709
17	1094	42.4	691	12	BG686343	BG686343 602638243
18	1072	41.6	974	12	BE989567	BE989567 601681494
19	1049	40.7	677	13	BI101409	BI101409 602887175
20	1034.5	40.1	959	12	BG250255	BG250255 602362379
21	1034	40.1	1062	12	BE895167	BE895167 601436096
22	1026	39.8	1111	10	BE311720	BE311720 60143309
c 23	1006	39.0	777	13	BI181203	BI181203 UNL-P-FN-
24	1003	38.9	624	10	AW476560	AW476560 ug77g04.Y
25	992	38.4	604	13	BI562276	BI562276 60355406
26	980.5	38.0	705	10	BG623528	BG623528 BB623528
27	979.5	38.0	700	9	AJ446923	AJ446923 AJ446923
28	976	37.8	634	13	BI067793	BI067793 Pgfln.pk0
29	955	37.0	614	13	BG913752	BG913752 602810906
30	945.5	36.6	742	13	BI18875	BI18875 603180927
31	945	36.6	614	12	BF087349	BF087349 QV2-HT054
32	921	35.7	789	13	BG913219	BG913219 602811885
33	911	35.3	546	10	BE257942	BE257942 601109776
34	911	35.3	594	9	AA726609	AA726609 vU93D06.f
35	907	35.2	843	12	BG425637	BG425637 602452939
36	889	34.5	625	9	AL585741	AL585741 AL585741
37	878.5	34.1	818	12	BE746542	BE746542 601580107
38	875	33.9	603	12	BG713339	BG713339 Pgfln.pk0
39	872	33.8	664	12	BG820485	BG820485 602782796
40	871	33.8	1159	14	BQ229819	BQ229819 AGENCOURT
41	865	33.5	782	13	BI253992	BI253992 602975130
42	862.5	33.4	527	9	AL118967	AL118967 DKFZP761E
43	838	32.5	635	9	AL783587	AL783587 AL783587
44	826.5	32.0	767	13	BI252061	BI252061 602952724
c 45	824	31.9	599	13	BI401530	BI401530 MI-P-CP0-

ALIGNMENTS

RESULT 1
BM560774 1036 bp mRNA linear EST 20-FEB-2002
LOCUS AGENCOURT 6566220 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5550695
DEFINITION 5' mRNA sequence.
ACCESSION BM560774
VERSION BM560774.1 GI:18805455
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1036)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

49 yo. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 2.1 kb, insert size range 1-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 036."

BASE COUNT 239 a 232 c 329 g 196 t
ORIGIN

Alignment Scores:
Pred. No.: 2,02e-139 Length: 1056
Score: 1465.00 Matches: 309
Percent Similarity: 95.69% Conservative: 2
Best Local Similarity: 95.08% Mismatches: 8
Query Match: 56.78% Indels: 7
DB: 14 Gaps: 0

US-09-830-144-4 (1-504) x BM906573 (1-1056)

Qy 1 MetAlaAlaGlnArgSerLeuLeuGlnSerGluGlnProSerTrpThrAspAsp 20
Db 12 ATGGCGGCGCAGAGGAGGAGCTTCTGCAGAGTGAGCAGCAAGCTGGACAGATGAC 71
Qy 21 LeuProLeuCysHisLeuSerGlyValGlySerAlaSerAsnArgSerTyrSerAlaAsp 40
Db 72 CTGCCTCTCTGCCACCTCTCTGGGGTTGGCTCAGCTCCAACCGCAGCTACTCTGCTGAT 131
Qy 41 GlyLysGlyThrGluSerHisProProGluAspSerTrpLeuLysPheArgSerGluAsn 60
Db 132 GGCAAGGCGACTGAGACCCACCGCCAGAGACAGCTGGCTCAAGTTCAGAGTGAGAAC 191
Qy 61 AsnCysPheLeuTyrGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnPheVal 80
Db 192 AACTGCTTCTGTATGGGCTCTTCAACGGCTATGATGGCAACCGAGTGACCAACTTCGTG 251
Qy 81 AlaGlnArgLeuSerAlaGluLeuLeuGlyGlnLeuAsnAlaGluHisAlaGluAla 100
Db 252 GCCCAGCGGCTGTCGAGAGCTCTCTGGCCAGCTGAATCCGAGCAGCCGAGGCGC 311
Qy 101 AspValArgArgValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGluSer 120
Db 312 GATGTGGGGCGTGTGCTGTCAGGCGCTTCGATGTGGTGAGAGGAGCTTCCTGGAGTCC 371
Qy 121 IleAspAlaLeuAlaGluLysAlaSerLeuGlnSerGlnLeuProGluGlyValPro 140
Db 372 ATTGACAGCCCTTGGCTGAGAAGGCAAGCCTCCAGTCGCAATTGCCAGAGGAGTCCCT 431
Qy 141 GlnHisGlnLeuProProGlnTyrGlnLysIleLeuGluArgLeuLysThrLeuGluArg 160
Db 432 CAGCACCAGCTGCCTCCTCAGTATCAGAAGATCCTTCAGAGACTCAAGACGTTAGAGAGG 491
Qy 161 GluIleSerGlyGlyAlaMetAlaValAlaValLeuLeuAsnAsnLysLeuTyrVal 180
Db 492 GAAATTCGGAGGGGCCATGCGCTGTGGCGGTCTCTTCTCAACAACCAAGCTCTACGTC 551
Qy 181 AlaAsnValGlyThrAsnArgAlaLeuLeuCysLysSerThrValAspGlyLeuGlnVal 200
Db 552 GCCAATGTCGGTCAAAACCGTCACATTTTATGCAAAATCGACAGTGATGGGTTGCAGGTG 611
Qy 201 ThrGlnLeuAsnValAspHisThrThrGluAsnGluAspGluLeuPheArgLeuSerGln 220
Db 612 ACACAGCTGAACGTGAGCACCACACACAGAGAACGAGGATGAGCTCTTCCTCTTCGCAG 671
Qy 221 LeuGlyLeuAspAlaGlyLysIleLysGlnValGlyIleLeuCysGlyGlnGluSerThr 240
Db 672 CTGGGCTTGATCTGGAAGATCAACAGAGTGGGATCATCTGTGGGAGAGAGACACC 731
Qy 241 ArgArgIleGlyAspTyrLysValLysTyrGlyTyrThrAspIleAspLeuLeuSerAla 260
Db 732 CGCGCGATCGGGATTACAG-GTTAAATATGGCTACACGACATTGACCTTCTCAGCGCT 790
Qy 261 AlaLysSerLysProIleAlaGluProGluIleHisGlyAlaGln-ProLeuAspGln 280

Db 791 GCCAAGTCCAAACCAATCATCGCAGAGCCGCAATCCCTGGGCGACAGCCCTGTGATGG 850
Qy 280 yVal-ThrGlyPheLeuValLeuMetSerGluGlyLeuTyrLysAla-LeuGluAlaAla 299
Db 851 GGTGAACGGCTTCTCTGGGCTGATGCCGAGGGGTGTACAAGGCCCTTAAGGCGAGCC 910
Qy 300 -HisGlyProGlyGlnAla-AsnGlnGluIleAlaAla-MetIleAspThrGluPheAla 318
Db 911 CAGGGGCCCGGGCAGGCCAAACAGAGAAATGTGTCACCAAGATTGACCTGAATTTGC 970
Qy 319 Lys 319
Db 971 CAA 973
RESULT 3
BO929862
LOCUS BO929862
DEFINITION BO929862 883 bp mRNA linear EST 20-AUG-2002
5', mRNA sequence.
ACCESSION BO929862
VERSION BO929862.1 GI:22344893
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 883)
NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LICM2648 row: c column: 19
High quality sequence stop: 705.
Location/Qualifiers
1..883
/organism="Homo sapiens"
/db xref="taxon:9606"
/clone="IMAGE:6460842"
/clone_lib="NIH MGC_101"
/tissue_type="epidermoid carcinoma, cell line"
/lab host="DH10B (phage-resistant)"
/notes="Organ: lung; Vector: pOTB7; Site 1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH MGC Library."

FEATURES
source

BASE COUNT 195 a -234 c 276 g 175 t 3 others
ORIGIN

Alignment Scores:
Pred. No.: 1.78e-130 Length: 883
Score: 1376.50 Matches: 274
Percent Similarity: 97.86% Conservative: 0
Best Local Similarity: 97.86% Mismatches: 4
Query Match: 53.35% Indels: 2
DB: 14 Gaps: 1

US-09-830-144-4 (1-504) x BO929862 (1-883)

Qy 7 SerLeuLeuGlnSerGluGlnProSerTrpThrAspLeuProLeuCysHisLeu 26
Db 2 AGCTTGCTGCAGAGTGAGCAGCAGCCAAAGCTGGACATGACCTGCTCTTGCACCTC 61

QY 27 SerGIValaIGlySerAlaSerAsnArgSerTyrSerAlaAspGlyLysGlyThrGluSer 46
 Db 62 TCTGGGGTGGCTCAGCCTCCCAACCGCAGCTACTGCTGATGGCAAGGAGCCTGAGAGC 121
 QY 47 HsSpProGluAspSerTyrLeuLysPheArgSerGluAsnAsnCysPheLeuTyrGly 66
 Db 122 CACCGCCGAGAGGACAGCTGGCTCAAGTTCAAGAGTAGAACAATTGCTTCTGTATGGG 181
 QY 67 ValPheAsnGlyTyrAspGlyAsnArgValThrAsnPheValAlaGlnArgLeuSerAla 86
 Db 182 GCTTCAACGGCTATGATGAGCAACGAGTGACCAACTTCTGTGGCCAGCGGCTGTCCGCA 241
 QY 87 GluLeuLeuLeuGlyGlnLeuAsnAlaGlnHisAlaGluAlaAspValArgValLeu 106
 Db 242 GAGCTCTCTGCTGGCCAGCTGAATGCCGAGCGCCGAGCGGAGTGGGGCTGTGGTG 301
 QY 107 LeuGlnAlaPheAspValValGluArgSerPheLeuGluSerIleAspAspAlaLeuAla 126
 Db 302 CTGCAAGCCTTGGATGTGGTGAGAGAGAGCTTCTGTGAGTCCATTGACGACCCCTGGCT 361
 QY 127 GluLysAlaSerLeuGlnSerGlnLeuProGluGlyValProGlnHisGlnLeuProPro 146
 Db 362 GAGAGGCAAGCCTCCAGTCGCAATGGCCAGAGGAGTCCCTCAGACACAGCTGCTCCT 421
 QY 147 GlnTyrGlnLysIleLeuGlnArgLeuLysThrLeuGlnArgGluIleSerGlyGlyAla 166
 Db 422 CAGTATCAGAGAGATCCTTGAAGACTCAAGACGTTAGAGAGGAAATTTCCGGAGGGGCC 481
 QY 167 MetAlaValAlaValLeuLeuAsnLysLeuTyrValAlaAsnValGlyThrAsn 186
 Db 482 ATGGCGGTGTGGGCTCTTCTCAACAACAAGCTTACGTGCGCAATGCGGTACCAAC 541
 QY 187 ArgAlaLeuLeuCysLysSerThrValAspGlyLeuGlnValThrGlnLeuAsnValAsp 206
 Db 542 CGGCACTTTTATGCAAAATCGACAGTGGATTGCAAGTGCAGTGACACAGCTGAACGTGAC 601
 QY 207 HisThrThrGluAsnGluAspGluLeuPheArgLeuSerGlnLeuGlyLeuAspAlaGly 226
 Db 602 CACACCACAGAGAACAGAGATGAGCTCTTCCGCTTTCCACAGTGGGCTGTGATCTGGA 661
 QY 227 LysIleLysGlnValGlyIleIleCysGlyGlnGluSerThrArgArgIleGlyAspTyr 246
 Db 662 AAGATCAAGAGAGGTGGATCATCTTGGGACAGAGAGACACCCGGCGATCCGGGATTC 721
 QY 247 LysValLysTyrGlyTyrThrAspIleAspLeuLeuSerAlaAlaLysSerLysProIle 266
 Db 722 AAGCTTAAATATGCTACACAGGACATTGACCTTCTCAGCGGCTGCCAAGTCCANACCAATC 781
 QY 267 IleAlaGluProGluIleHis-GlyAlaGlnProLeuAsp---GlyValThrGlyPhe 284
 Db 782 ATCGCAGAGCCAGAAATCCATGGGGCAGAGCCCGCTGGAGATGGGGTGACGGGCTTT 839
 RESULT 4
 BQ922049 863 bp mRNA linear EST 20-AUG-2002
 LOCUS BQ922049
 DEFINITION AGENCOURT 8923996 NIH_MGC_94 Mus musculus cDNA clone IMAGE:6467676
 5', mRNA sequence.
 ACCESSION BQ922049
 VERSION BQ922049.1 GI:22337080
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 863)
 NIH-MGC <http://mgs.nci.nih.gov/>
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: The Cepko Laboratory
 cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LMNL at:
<http://image.llnl.gov>
 Plate: LLM13993 row: p column: 13
 High quality sequence stop: 654.
 Location/Qualifiers
 1..863
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="IMAGE:6467676"
 /clone_1ib="NIH_MGC_94"
 /tissue_type="retina"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: pCMV-SpOrf6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 3.3 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

BASE COUNT 193 a 234 c 264 g 171 t 1 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 5,59e-128 Length: 863
 Score: 1352.00 Matches: 276
 Percent Similarity: 97.59% Conservative: 7
 Best Local Similarity: 95.17% Mismatches: 2
 Query Match: 52,40% Indels: 6
 DB: 14 Gaps: 2

US-09-830-144-4 (1-504) X BQ922049 (1-863)
 QY 107 LeuGlnAlaPheAspValValGluArgSerPheLeuGluSerIleAspAspAlaLeuAla 126
 Db 3 CTTCAGGCTTTGATGTGTGGAGAGAGCTTCTGAGTCTATCGATGATGCTTACCT 62
 QY 127 GluLysAlaSerLeuGlnSerGlnLeuProGluGlyValProGlnHisGlnLeuProPro 146
 Db 63 GAGAAACCAAGCTCCAGTCCAGCTGCGCGAGAGGTGTGCCACAGACCAAGCTGCACCT 122
 QY 147 GlnTyrGlnLysIleLeuGlnArgLeuLysThrLeuGlnArgGluIleSerGlyGlyAla 166
 Db 123 CAGTATCAGAGAGATCCTTGAAGACTCAAGCACTGAGAGGGAGATTTCCGGAGGAGGCC 182
 QY 167 MetAlaValAlaValLeuLeuAsnLysLeuTyrValAlaAsnValGlyThrAsn 186
 Db 183 ATGGCTGTCTGGGCAGTCTTCTCAACAGCAAGCTTACGTTGCCAATGCGGTACGAAC 242
 QY 187 ArgAlaLeuLeuCysLysSerThrValAspGlyLeuGlnValThrGlnLeuAsnValAsp 206
 Db 243 AGGCCCTCTGTGCAAAATCTACAGTGGATGGATTACAGGTTACACAGCTTAACATGAC 302
 QY 207 HisThrThrGluAsnGluAspGluLeuPheArgLeuSerGlnLeuGlyLeuAspAlaGly 226
 Db 303 CACACCAACGAGACGAGAGAGCTCTTGGCTTTCGCAATGGGTTTACAGCGAGGG 362
 QY 227 LysIleLysGlnValGlyIleIleCysGlyGlnGluSerThrArgArgIleGlyAspTyr 246
 Db 363 AAGATCAAGCAGATGGGCTCATCTGTGACAGAGAGCACAGCGCGGATGGGGATTAC 422
 QY 247 LysValLysTyrGlyTyrThrAspIleAspLeuLeuSerAlaAlaLysSerLysProIle 266
 Db 423 AAGGTCAAAATATGGCTACCGCAGCATTGACTGTCAAGCTCCCAAGTCCAAACCATC 482
 QY 267 IleAlaGluProGluIleHisGlyAlaGlnProLeuAspValValThrGlyPheLeuVal 286
 Db 483 ATCGCAGCGCGAAATTCATGGGCAAGCTTGTGATGGCTGTGACGGGCTTCTGTG 542
 QY 287 LeuMetSerGluGlyLeuTyrLysAlaLeuGlnAlaAlaHisGlyProGlyGlnAlaAsn 306
 Db 543 CTGATGTAGAGGGGCTGTCAAGAGCCCTGAGGCAACCATGAGGCTTGGGAGGCCAAC 602

JOURNAL COMMENT

Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at:
 http://image.llnl.gov
 Plate: LLM11627 row: 0 column: 03
 High quality sequence stop: 841.
 Location/Qualifiers

FEATURES source

1..847
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5248730"
 /clone_lib="NIH_MGC_121"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 3 fetal brains, female age 20 weeks, female age 24 weeks, and male age 26 weeks. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb. Insert size range 0.7-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 017. Note: this is a NIH MGC Library."
 BASE COUNT 195 a 229 c 264 g 159 t
 ORIGIN

Alignment Scores:

Prod. No.: 1,17e-120 Length: 847
 Score: 1280.50 Matches: 269
 Percent Similarity: 95.42% Conservative: 2
 Best Local Similarity: 94.72% Mismatches: 6
 Query Match: 49.63% Indels: 7
 DB: 13 Gaps: 1

US-09-830-144-4 (1-504) x BI915996 (1-847)

QY 4 GlnArgSerLeuLeuGlnGlnProSerTrpThrAspLeuProLeu 23
 Db 3 CAGAGAGAGAGCTTGGCTGCAGAGTGAAGCAGCAAGCTGACATGACCTGCTTC 62
 QY 24 CysHisLeuSerGlyValGlySerAlaSerAsnArgSerTyrSerAlaAspGlyLysGly 43
 Db 63 TGCACCTCTCTGGGAGGCTCAGCTCCACCGAGCTACTCTCTGATGGCAAGGC 122
 QY 44 ThrGluSerHisProProGluAspSerTrpLeuLysPheArgSerGluAsnAsnCysPhe 63
 Db 123 ACTGAGAGCAACCCGCCAGAGACAGCTGCTCAAGTTCAGAGTGAAGCAACTCTTC 182
 QY 64 LeuTyrGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnPheValAlaGlnArg 83
 Db 183 CTGTATGGGGTCTTCAACGGCTATGATGGCAACCGAGTCCAACTTCGTGGCCAGCGG 242
 QY 84 LeuSerAlaGluLeuLeuGlnGlnLeuAsnAlaGlnHisAlaGluAlaAspValArg 103
 Db 243 CTGTCGCCAGAGCTCTGGCTGGGCGACACTGAATGCCAGACCCCGAGCCGATGTCGG 302
 QY 104 ArgValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGlnSerIleAspAsp 123
 Db 303 CGTGTGCTGTCAGAGCTTCGATGTGTGGAGAGGAGCTTCCTGGAGTCCATTGACGAC 362
 QY 124 AlaLeuAlaGluLysAlaSerLeuGlnSerGlnLeuProGluGlyValProGlnHisGln 143
 Db 363 GCCTTGCTGAGAAAGCAAGCTCCAGTCCCAATTGGCAAGGAGATGCTCTCAGCACGAG 422
 QY 144 LeuProProGlnTyrGlnLysIleLeuGlnArgLeuLysThrLeuGlnArgGluLysSer 163
 Db 423 CTGCTCTCTCAATCAAGATCTTGAAGACTCAAGACGTTAAGAGAGGAAATTTG 482

QY 164 GlyGlyAlaMetAlaValAlaValLeuLeuAsnLysLeuTyrValAlaAsnVal 183
 Db 483 GAGAGGGCCATGAGCCGTGTGGGGGTCTTCTTCACACAGCTTCATCGCCAAATGTC 542
 QY 184 GlyThrAsnArgAlaLeuLeuCysLysSerThrValAspGlyLeuGlnAlaThrGlnLeu 203
 Db 543 GGTACAAACCGTGACTTTATGCAATTCACAGTGAATGGTGTGAGGTGACACAGCTG 602
 QY 204 AsnValAspHisThrThrGluAsnGluAspGluLeuPheArgLeuSerGlnLeuGlyLeu 223
 Db 603 AACGTGACCAACACACAGAGAAAGAGATGAGCTCTTCGCTTTCCAGCTGGCTTG 662
 QY 224 AspAlaGlyLysIleLysGlnValGly-IleIleCys-GlyGlnGlu-SerThrArgArg 242
 Db 663 GATCTGGAAGAATCAAGCGAGTGGCGGATCATCTGTGGGCGAGAGACCCCGCG 722
 QY 243 Ile-GlyAspTyrLysValLysTyrGlyTyrThr-AspIleAspLeuLeu-SerAlaAla 261
 Db 723 ATCGGGGATTTACAGGTTCAATATGGCTTACCCGACATTGACCTTCTCCAGCGCTGCC 782
 QY 262 LysSerLysProIleIleAlaGluPro--GluIleHisGlyAlaGlnProLeuAspGly 280
 Db 783 AAGGTCAACCAATCATCGACAGAGCCAGAAATCCATGGCGCACAGCGCTGATGG 842

RESULT 7
 B0644850
 LOCUS
 DEFINITION
 AGENCOURT 8287994 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:5298124
 5', mRNA sequence.
 ACCESSION
 B0644850
 VERSION
 B0644850.1 GI:21769022
 KEYWORDS
 EST.
 SOURCE
 human.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 892)
 NIH-MGC http://mgs.nci.nih.gov/.
 TITLE
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL
 Unpublished (1999)
 CONTACT
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: CGAP (Stanford)
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at:
 http://image.llnl.gov
 Plate: LLM2508 row: k column: 21
 High quality sequence stop: 637.
 Location/Qualifiers

FEATURES source

1..892

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6298124"
 /clone_lib="NIH_MGC_100"
 /tissue_type="hepatocellular carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: liver; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
 BASE COUNT 202 a 255 c 273 g 160 t 2 others
 ORIGIN

Alignment Scores:

Pred. No.:	1.43e-120	Length:	892
Score:	1280.00	Matches:	274
Percent Similarity:	94.00%	Conservative:	8
Best Local Similarity:	91.33%	Mismatches:	8
Query Match:	49.61%	Indels:	10
DB:	14	Gaps:	3
US-09-830-144-4 (1-504) x BQ644850 (1-892)			
Qy	123	AspAlaLeuAlaGluLysAlaSerLeuGlnSerGlnLeuProGluGlyValProGlnHis	142
Db	3	GACGCTTGGCTGAGAGGCAAGCTCCAGTCGCAATTGCCAGAGGAGTCCCTCAGCAC	62
Qy	143	GlnLeuProGlnTyrGlnLysIleLeuGluArgLeuLysThrLeuGluArgGluIle	162
Db	63	CAGTGCCTCCTCAGTATCAGAAGATCCTTGAGAGACTCAAGACGTTAGAGAGGAATT	122
Qy	163	SerGlyGlyAlaMetAlaValAlaValLeuLeuLeuAsnLysLeuTyrValAlaAsn	182
Db	123	TCGGAGGGGCAATGGCGTTGTGGCGTCTCTCAACAACAAGCTCTACGTCGCAAT	182
Qy	183	ValGlyThrAsnArgAlaLeuLeuCysLysSerThrValAspGlyLeuGlnValThrGln	202
Db	183	GTCGGTACAAACCGTGACITTTATGCAATCGACATGGATGGTTGCAGGTGACACAG	242
Qy	203	LeuAsnValAspHisThrThrGluAsnGluAspGluLeuPheArgLeuSerGlnLeuGly	222
Db	243	CTGAACGTGGACACACACACAGAGAGAGATGAGCTCTTCGCTCTTCGACGTGGGC	302
Qy	223	LeuAspAlaGlyLysIleLysGlnValGlyIleIleCysGlyGlnGluSerThrArgArg	242
Db	303	TTGGATGCTGAAAGATCAACAGAGTGGGATCATCTGTGGCAGGAGACACCGCGCG	362
Qy	243	IleGlyAspTyrLysValLysTyrGlyTyrThrAspIleAspLeuLeuSerAlaAlaLys	262
Db	363	ATCGGGATTACAAGTTAAATATGGCTACAGGACATTGACCTTTCACCGCTGCCAAG	422
Qy	263	SerLysProIleIleAlaGluProGluIleHisGlyAlaGlnProLeuAspGlyValThr	282
Db	423	TCCAACAACATCATCGCAGAGCCAGAAATCCATGGGCGCAGCCGCTGGATGGGTGACG	482
Qy	283	GlyPheLeuValLeuMetSerGluGlyLeuTyrLysAlaLeuGluAlaHisGlyPro	302
Db	483	GGCTTCTTGGTGCTGATGTCGGAGGGGTGTACAAGGCCCTAGAGGCAGCCCTTGGCCT	542
Qy	303	GlyGlnAlaAsnGlnGluIleAlaAlaMetIleAspThrGluPheAlaLysGlnThrSer	322
Db	543	GGGCAGCCCAACCAAGAGATTGTCGATGATTGACACTGAGTTTCCCAAGCAGACCTCC	602
Qy	323	LeuAspAlaValAlaGlnAlaValAspArgValLysArgIleHisSerAspThrPhe	342
Db	603	CTGGACGACGTGGCCCAAGGCGCTCGTGGACCGGTGAAGCGCATCCACAGCGACCTTC	662
Qy	343	AlaSer---GlyGlyGluArgAlaArg-PheCys-ProArgHisGluAspMetThr-Leu	360
Db	663	GCAGTGGGGGGGGAACCGTGCAGGNNCTGCCCGCCGCCACGAGGACATGACCCCTTG	722
Qy	361	LeuValArgAsn-PheGlyTyrProLeu-GlyGluMetSerGlnProThrProSerProA	380
Db	723	CTAGTGAGGAACCTTTGGCTACCGCTGGGGCGGAATATGACCCAGCCACACCGAGCCAG	782
Qy	380	laProAlaAla-GlyGlyArgValTyrProValSerValProTyrSerSerAla-----	397
Db	783	CCCCAGTGCAGGAGGACGAGTGTACCCCTGTGGTCTGTGGCCATACCTCCAGCGCCCCA	842
Qy	398	GlnSerThrSerLysThrSerValThrLeuSer---LeuValMetPro	412
Db	843	GAGGCACAGCAAGAACCGGTGGACCTTCTCCCTTGTTCATGCC	890
RESULT 8			
BQ645086			
LOCUS			
DEFINITION			
AGENCOURT_8493044 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6300450			
991 bp mRNA linear EST 15-JUL-2002			

5', mRNA sequence.			
ACCESSION	BQ645086		
VERSION	BQ645086.1	GI:21769258	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 991)		
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: CGAP (Stanford) cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLCM2514 row: 1 column: 19 High quality sequence stop: 534. Location/Qualifiers		
FEATURES	1..991		
source	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:6300450"		
	/clone_lib="NIH_MGC_100"		
	/tissue_type="hepatocellular carcinoma, cell line"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: liver; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."		
BASE COUNT	236 a	292 c	292 g
ORIGIN			166 t
			5 others
Alignment Scores:			
Pred. No.:	4.15e-119	Length:	991
Score:	1266.50	Matches:	286
Percent Similarity:	86.79%	Conservative:	3
Best Local Similarity:	85.89%	Mismatches:	30
Query Match:	49.09%	Indels:	14
DB:	14	Gaps:	3
US-09-830-144-4 (1-504) x BQ645086 (1-991)			
Qy	123	AspAlaLeuAlaGluLysAlaSerLeuGlnSerGlnLeuProGluGlyValProGlnHis	142
Db	3	GACGCTTGGCTGAGAGGCAAGCTCCAGTCGCAATTGCCAGAGGAGTCCCTCAGCAC	62
Qy	143	GlnLeuProGlnTyrGlnLysIleLeuGluArgLeuLysThrLeuGluArgGluIle	162
Db	63	CAGTGCCTCCTCAGTATCAGAAGATCCTTGAGAGACTCAAGACGTTAGAGAGGAATT	122
Qy	163	SerGlyGlyAlaMetAlaValAlaValLeuLeuLeuAsnLysLeuTyrValAlaAsn	182
Db	123	TCGGAGGGGCGCATGGCGTTGTGGCGTCTCTCAACAACAAGCTCTACGTCGCAAT	182
Qy	183	ValGlyThrAsnArgAlaLeuLeuCysLysSerThrValAspGlyLeuGlnValThrGln	202
Db	183	GTCGGTACAAACCGTGACITTTATGCAATCGACATGGATGGTTGCAGGTGACACAG	242
Qy	203	LeuAsnValAspHisThrThrGluAsnGluAspGluLeuPheArgLeuSerGlnLeuGly	222
Db	243	CTGAACGTGGACACACACACAGAGAGATGAGCTCTTCGCTCTTCGACGTGGGC	302

QY 223 LeuAspAlaGlyLysIleuSgInValIylIleIcySgIyngInuSerThrArg 242
 Db 303 TTGGATGCTGGAAAAGATCAAGACGGTGGGATCATCTGGGACAGAGACCCGGGG 362
 QY 243 TleGlyAspTyrLysValIleYrIyThrAspIleAspLeuSerAlaIals 262
 Db 363 ATCGGGGATTACAGAGTTAAATATGCTACAGACGACATTGACCTTCTAGCGCTGCCAAG 422
 QY 263 SerLysProIleIleAlaGluProGluIleHisGlyAlaGlnProLeuAspGlyValThr 282
 Db 423 TCCAAACCATATATCGCAGAGCCAGAAATCCATGGGGCAGACCGCTGATGGGGTGACG 482
 QY 283 GlyPheLeuValLeuMetSerGluGlyLeuTyrLysAlaLeuGluAlaIleHisGlyPro 302
 Db 483 GGCTCTTGCTGCTGATGTCGAGGGGTTGTACAAAGGCCCTTAGAGGACCCATGGGGCT 542
 QY 303 GlyGlnAlaAsnGlnGluIleAlaIleMetIleAspThrGluPheAlaLysGlnThrSer 322
 Db 543 GGGCAGGCCAACCCAGAGATTGCTCGATGATTGACACTGAGTTTCCAGACGACCTTC 602
 QY 322 LeuAspAlaValAlaGlnAlaValAlaAspArgValLysArgIleHisSerAspThrPhe 342
 Db 603 CTGGACGCGCTGGCCAGGCCCTCGTGCAGCCGGTGGAAGCGCATCCACAGGACACCTTC 662
 QY 343 AlaSerGly--GlyGluArgAlaArgPheCysProAlaGHisGluAspMetThr--LeuLe 361
 Db 663 NNCAGGGGTGGGGAGCGGTGCCANNGTTCTGCCCGGACAGAACATGACCTCTGCT 722
 QY 361 ValAlaArg--AsnMetGlyTyr--ProLeuGlyGluMetSerGlnProThrProSerProAla 380
 Db 723 AGTGAAGGAACCTTGGCTACCCCTCGCGCAATGAGCACCCACACCGAGCGCCCTCC 782
 QY 381 ProAlaAla--GlyGlyArgValTyrPro--ValSerValProTyrSer--SerAlaGlns 399
 Db 783 CCANCTTGCAGAGAGACCAATGTACCTCGGGGTCTGTGCATCTCCAGGGGCCCAA 842
 QY 399 eThr--SerLysThrSer--ValThrLeuSerLeuValMetProSerGlnGlyMet-- 417
 Db 843 GCACCCACCAAGACCGAGCGGGACCTCTCCCTTGCCATGCTCCCGGCGCAAAAAG 902
 QY 418 -ValAsnGlyAlaHis--SerAlaSerThrLeuAspGluAlaThrProThr--LeuThrA 436
 Db 903 GCCACAGGGGCTCAAAAGGGGCTTCCCCCGGAAAGAAAACCCCCCCTTACCA 962
 QY 436 sngInSerProThrLeu 441
 Db 963 AACAAAGGCCGACACTT 979
 RESULT 9
 BQ42448 784 bp mRNA linear EST 29-MAY-2002
 LOCUS UI-M-EV0-dxe-g-20-0-UI.r1 NIH BMAP_EV0 Mus musculus cDNA clone
 DEFINITION IMAGE:5707051.5', mRNA sequence.
 ACCESSION BQ42448
 VERSION BQ42448.1 GI:21245560
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 784)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 CDNA library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)
 Seq primer: pyx-5,
 Location/Qualifiers
 source
 1..784
 /organism="Mus musculus"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:5707051"
 /clone_1ib="NIH BMAP_EV0"
 /tissue_type="whole brain"
 /dev_stage="embryo 15.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /note="Organ: brain; Vector: pyx-asc; Site_1: EcoR I;
 Site_2: Not I; The library was constructed according to
 Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured mRNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with an
 oligo-dT primer containing a Not I site. Double stranded
 cDNA was size selected according to mRNA size fraction,
 ligated with EcoR I adaptor, digested with Not I, and then
 cloned directionally into pyx-asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is GTCGCTGAGAA. This library was created for the
 University of Iowa Mouse Brain Molecular Anatomy Project
 (BMAP): 'Gene Discovery in the Developing Mouse Nervous
 System', supported by National Institutes of Mental Health
 (NIMH), Hemm Chin, Ph.D., program coordinator."

BASE COUNT 180 a 228 c 223 g 150 t 3 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 9,64e-114 Length: 784
 Score: 1212.50 Matches: 245
 Percent Similarity: 96.18% Conservative: 8
 Best Local Similarity: 93.51% Mismatches: 7
 Query Match: 47,00% Indels: 4
 DB: Gaps: 1

US-09-830-144-4 (1-504) x BQ42448 (1-784)
 QY 188 AlaleuencysylsserThrValaSpGlyLeuGlnValThrGlnLeuAsnValaSpHis 207
 Db 3 GCCCTTCTGCAAAATTACAGTGGATGGTTACAGGTTACACGTTAAACATGACGAC 62
 QY 208 ThrThrGluAsnGluAspGluLeuPheArgLeuSerGlnLeuGlyLeuAspAlaGlyLys 227
 Db 63 ACCACGAGAACGAGACGAGCAGCTCTTCGGCTTCCCACTGGGTTAGCGCAGGGAAG 122
 QY 228 TleYsgInValGlyIleIleIcySgIyngInuSerThrArgArgIleGlyAspTyrLys 247
 Db 123 ATCAAGCAGATGGCGCTGCTGTGACAGAGACCCAGGGGATGGGGATTTCAAG 182
 QY 248 ValLysTyrGlyTyrThrAspIleAspLeuSerAlaIalsSerLysProIleIle 267
 Db 183 GTCAAAATATGGCTACACCGACATTGACCTCTCAGCGCTCCAAAGTCCAAACCATCATC 242
 QY 268 AlaGluProGluIleHisGlyAlaGlnProLeuAspGlyValThrGlyPheLeuValLeu 287
 Db 243 GCAAGCGCGAAATTCATGGGGCAGACCTCTGATGCGCTGACGGGCTCTCGTGCTG 302
 QY 288 MetSerGluGlyLeuTyrLysAlaLeuGluAlaIleHisGlyProGlyGlnAlaGln 307
 Db 303 ATGTCAAGGGGCTGTCAAGGGCTTGAGAGCCATGCGCTGGCGGCGCAACGAG 362
 QY 308 GluIleAlaIleMetIleAspThrGluPheAlaLysGlnThrSerLeuAspAlaValAla 327
 Db 363 GAGATTGCCCGGATGATTGACCGCAATTGGCCAGAGACCTCCCTGATGGGCTTCCC 422
 QY 328 GlnAlaValAlaAspArgValLysArgIleHisSerAspThrPheAlaSerGlyGlyGlu 347
 Db 423 CAGCGTGTGTGAGACCGGTGTAAGCGGATCCACAGTACACCTTTGCGCATGTGGGGAG 482


```

VERSION      BI181306.1  GI:14655715
KEYWORDS     EST.
SOURCE       Pig.
ORGANISM     Sus scrofa
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE    1 (bases 1 to 815)
AUTHORS      Caetano, A.R., Johnson, R.K. and Pomp, D.
TITLE        Generation and sequence characterization of a normalized cDNA
              library from swine ovarian follicles
              Unpublished (2001)
JOURNAL      Contact: Pomp, D
              Department of Animal Science
              University of Nebraska, Lincoln
              Lincoln, NE 68583-0908, USA
              Tel: 402 472 6416
              Fax: 402 472 6362
              Email: dpomp@unl.edu
COMMENT      Oligo-dT track not found, Not 1 site shown in beginning of sequence
              is likely internal to the message. The following repetitive
              elements were found in this cDNA sequence: 62-112,
              >GC rich#low complexity
              Seq primer: M13 -29
              POLYA=No.

FEATURES
    source
        1..815
            /organism="Sus scrofa"
            /strain="University of Nebraska, Lincoln Swine Selection
            lines"
            /db_xref="taxon:9823"
            /clone="UNL-P-FN-au-g-01-0-UNL"
            /clone_1ib="UNL-P-FN"
            /dev_stage="ADULT"
            /lab_host="DH10B (Life Technologies)"
            /note="Vector: pUT73D-pac (pharmacia) with a modified
            polylinker. Site_1: Not I; Site_2: Eco RI; The UNL-P-FN
            library is a normalized library representing porcine
            ovarian follicles, ranging between 2.0 to 10.0 mm in
            diameter, collected during 7 days of the follicular phase
            of the pig estrous cycle. This library was derived from
            the library UNL-P-F2. The tag is a string of 5-6
            nucleotides present between the Not I site and the
            oligo-dT track. The library was constructed as described
            by Bonaldo, Lennon and Soares, Genome Research 6: 791-806
            , 1996.
            TAG_SEQ=None found"
BASE COUNT   109 a      266 c      306 g      134 t
ORIGIN
Alignment Scores:
Pred. No.:    6,05e-109      Length:      815
Score:        1166.00        Matches:    227
Percent Similarity: 98.72%    Conservative: 4
Best Local Similarity: 97.01% Mismatches:    2
Query Match:  45.19%         Indels:      1
DB:           13            Gaps:        0
US-09-830-144-4 (1-504) x BI181306 (1-815)
QY 272 ILHISGLYALAGLNPProleuAspGlyValThrGlyPheLeuValLeuMetSerGluGly 291
DB 814 ATCCACGGTGCACAGCCCGCGGAGCGGGTGCAGCGCTTCCTGAGTGCAGTCCGAGGG 755
QY 292 LeuTyrlLeuAlaLeuGluAlaAlaHisGlyProGlyGlnAlaA-AspGlnGluIleAla 311
DB 754 CTATACAGGCCCTCGAGAGCGAGCCCGGCGCTGGCGCAGGCCAAACAGAGATCGCAGC 695
QY 311 aMetIleAspThrGluPheAlaLysGlnThrSerLeuAspAlaValAlaGlnAlaValVa 331
DB 694 CATGATCGACACAGAGATTGCCAAGCAGACTCCCTGGATGCGCTGGCCAGGCCCTGTGT 635
QY 331 laapaGValLysArgIleHisSerAspThrPheAlaSerGlyGluArgAlaArgph 351

```

```

DB 634 GGACCGGGTGAAGCGATCCACAGTGCACACTTCGCCACCGGGGGAGCGTCCAGTT 575
QY 351 eCySPARhISGLIAspMetThrLeuLeuValArgAspMetGlyTyrlProLeuGlyG 371
DB 574 CTGCCACGGCAGAGAGACATGACCTGCTGGTGGCAGACTTCCTTCCTGGCAGCA 515
QY 371 uMetSerGlnProThrProSerProAlaProAlaHisGlyGlyArgValTyrlProValSe 391
DB 514 GCTAGGACACCCACAGCGCGGCCACCCCGAGCGGAGAGAGAGCGGTATACCCCTGTTC 455
QY 391 rValProTySerSerAlaGlnSerThrSerLysThrSerValThrLeuSerLeuValMe 411
DB 454 TGTGCCACTCTCAGAGCCCGCCAGACACACAGACAGACAGTACCTGTCTCTCTCAT 395
QY 411 tProSerGlnGlyGlnMetValAsnGlyAlaHisSerAlaSerThrLeuAspGluAlaTh 431
DB 394 GCCCTCCAGGCGCAGCTGTCAACGCGGGCCACAGTGTCTTCACCTCGAGCAACGCCAC 335
QY 431 rProThrLeuThrAsnGlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrG 451
DB 334 GCCACCTTCACCAACCAAGAGCCCGACCTGACCTTCAGTCCACCAACCCACCA 275
QY 451 nSerSerSerSerSerSerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuPr 471
DB 274 GAGCAGACGCTCCAGCTCCACGGGGGCTCTTCCTCCCGGCCACACCTCGCTCCC 215
QY 471 oProGlyGluAspGlyArgValGluProTyrlValAspPheAlaGluPheTyrlArgLeuTr 491
DB 214 GCCCGGAGAGATGCGCGCGCTGAGCCCTGAGCTTCGAGACTTCGCGAGATTACCGCTGTG 155
QY 491 pSerValAspHisGlyGlnGlnSerValValThrAlaPro 504
DB 154 GAGGTTGACACCGCGCAGAGAGTGTGGAGCGGGCGCG 115
RESULT 12
BG424017 910 bp mRNA linear EST 14-MAR-2001
BG424017
LOCUS      602447479F1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:4586298 5',
DEFINITION mRNA sequence.
ACCESSION BG424017 GI:13330523
VERSION   BG424017.1
KEYWORDS  EST.
SOURCE    human.
ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
            1 (bases 1 to 910)
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
JOURNAL   Contact: Robert Strausberg, Ph.D.
            Email: cga@bs-r@mail.nih.gov
            Tissue Procurement: DCTD/DRP
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLM314 row: m column: 19
            High quality sequence stop: 692.
FEATURES
    source
        1..910
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:4586298"
            /clone_1ib="NIH_MGC_14"
            /tissue_type="renal cell adenocarcinoma"
            /lab_host="DH10B (phage-resistant)"
            /note="Organ: Kidney; Vector: pOTB7; Site_1: XhoI; Site_2:
            EcoRI; cDNA made by oligo-dT priming. Directionally
            cloned into EcoRI/XhoI sites using the following 5'
            adaptor: GGCAAGAG(G). Size-selected >500bp for average

```


insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 197 a 241 c 297 g 175 t

ORIGIN

Alignment Scores:
Pred. No.: 8,21e-109 Length: 910
Score: 1165.50 Matches: 256
Percent Similarity: 92.81% Conservatve: 2
Best Local Similarity: 92.09% Mismatches: 15
Query Match: 45.17% Indels: 8
DB: 12 Gaps: 1

US-09-830-144-4 (1-504) x BG424017 (1-910)

Qy 1 MetAlaAlaGlnArgSerLeuLeuGlnSerGluGlnGlnProSerTrpThrAspAsp 20
Db 9 ATGGCGGCGCAGAGAGAGGCTTCTCAGAGTGCAGAGCAGCAGCCAACTGGACAGATGAC 68
Qy 21 LeuProLeuCysHisLeuSerGlyValGlySerAlaSerAsnArgSerTyrSerAlaAsp 40
Db 69 CTGCTCTCTCCACCTCTCTGGGGTGGCTCAGCTCCACCGCAGCTACTCTGCTGAT 128
Qy 41 GlyLysGlyThrGluSerHisProGluAspSerTrpLeuLysPheArgSerGluAsn 60
Db 129 GGCAGGGCACTGAGAGCCACCCGAGAGCAGCTGGCTCAAGTTCAGAGTGGAGAAC 188
Qy 61 AsnCysPheLeuTyrGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnPheVal 80
Db 189 AACTGCTTCCTGTATGGGGTCTTCAACGGCTATGATGGCAACCGAGTGACCAATTCGTG 248
Qy 81 AlaGlnArgLeuSerAlaGluLeuLeuGlyGlnLeuAsnAlaGluHisAlaGluAla 100
Db 249 CCCAGCGCTGTTCGCAGAGCTCTGCTGGCCAGCTGAATGCCGAGCAGCCGAGGCC 308
Qy 101 AspValArgArgValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGluSer 120
Db 309 GATGTGGCGCTGTGCTCAGCGCTTCGATGTGTGGAGAGAGCTTCCTGGAGTCC 368
Qy 121 IleAspAspAlaLeuAlaGluLysAlaSerLeuGlnSerGlnLeuProGluGlyValPro 140
Db 369 ATTACGACGCTTGGCTGAGAAGCAAGCTCCAGTTCGCAATTCGACAGAGGAGTCCCT 428
Qy 141 GlnHisGlnLeuProGlnTyrGlnLysIleLeuGluArgLeuLysThrLeuGluArg 160
Db 429 CAGCACCACTGCTCTCAGTATCAGAAGATCCTTGAGAGACTCAAGAGCTTAGAGAGG 488
Qy 161 GluIleSerGlyGlyAlaMetAlaValAlaValLeuLeuAsnAsnLysLeuTyrVal 180
Db 489 GAAATTCGGGAGGGCCATGGCCGTGTGGCGTCTCTCAACAACAGCTCTAGCTC 548
Qy 181 AlaAsnValGlyThrAsnArgAlaLeuLeuCysLysSerThrValAspGlyLeuGlnVal 200
Db 549 GC-AATGTTCGTCACAAACCGTGCACTTTATGCAATCGACAGTGGATGGGTTCGAGGTG 607
Qy 201 ThrGlnLeuAsnVal-AspHisThrThrGluAsnGluAspGluLeuPhe-ArgLeu-Ser 219
Db 608 ACACAGCTGAACGTGGACCAACACACAGAGAACGAGGATGAGCTCTTCCCGTCTTTTCG 667
Qy 220 GlnLeuGlyLeuAspAlaGlyLysIleLysGlnValGlyIleIleCys-GlyGlnGluSe 239
Db 668 CAGCTGGTTCGGATGCTGGAAAGATACCGCAGGTGGGGATCATCTGTGGGGCAGGAG 727
Qy 239 rThrArgGileGlyAspTyrLysValLysTyrGlyTyrThrAspIleAspLeuSe 259
Db 728 CACCGGGGGATCGGGATTA-ACCGTTAACTATGGGTTCACGGA-ATTGACTTCTTAGG 785
Qy 259 rAlaAlaLysSerLysProIleIleAlaGluProGluIleHisGly 274
Db 786 GGGTGCC---AGTCAACAACATCATCGAGAGGCCCAAACTCTGGGG 828

RESULT 13
BQ179703
LOCUS

DEFINITION

ACCESSION
VERSION
KEYWORDS
SOURCE

ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES
source

Location/Qualifiers
1..714

/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5703067"
/clone_lib="NIH BMAP_EWO"
/tissue_type="whole brain"
/dev_stage="embryo 15.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/note="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was size selected according to mRNA size fraction,
ligated with EcoR I adaptor, digested with Not I, and then
cloned directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
, is GTGCGTGGAA. This library was created for the
University of Iowa Mouse Brain Molecular Anatomy Project
(BMAP): 'Gene Discovery in the Developing Mouse Nervous
System', supported by National Institutes of Mental Health
(NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT 167 a 186 c 213 g 146 t

ORIGIN

Alignment Scores:
Pred. No.: 2,12e-107 Length: 714
Score: 1150.00 Matches: 228
Percent Similarity: 97.90% Conservatve: 5
Best Local Similarity: 95.80% Mismatches: 4
Query Match: 44.57% Indels: 2
DB: 14 Gaps: 0

US-09-830-144-4 (1-504) x BQ179703 (1-714)

Qy 9 LeuGlnSerGluGlnGlnProSerTrpThrAspLeuProLeuCysHisLeuSerGly 28
Db 2 CTCGAGGTGACGACGCCAAGCTGGACAGATGACCTGCCACTCTGTCACTCTCTCGA 61
Qy 29 ValGlySerAlaSerAsnArgSerTyrSerAlaAspGlyLysGlyThrGluSerHisPro 48
Db 62 GTTGTGTCAGCGTCCAAACCGCAGCTACTCTGCTGATGCAAGGGCACTGAGAGCCACCT 121

BQ179703 714 bp mRNA linear EST 30-APR-2002
UI-M-EWO-Bwu-a-20-0-UI_r1 NIH BMAP_EWO Mus musculus cDNA clone
IMAGE:5703067 5', mRNA sequence.

BQ179703
BQ179703.1 GI:20355195
EST.

house mouse.
Mus musculus.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 714)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: c9apbs-remail.nih.gov

Tissue Procurement: Dr. James Lin, Univeristy of Iowa

cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LUNL at:

http://image.llnl.gov

This clone was contributed by the Brain Molecular Anatomy Project

(BMAP)

Seq primer: pYX-5.

QY 49 ProGluAspSerTrpLeuLysPheArgSerGluAsnAsnCysPheLeuTyrGlyValPhe 68
DB 122 CCGAGAGCAACTGCTTAAGTTCAGAAAGTGAATAACTGCTTCCTGACCGGGCTCTTC 181
QY 69 AasnGlyTyrAspGlyAsnArgValThrAsnPheValAlaGlnArgLeuSerAlaGluLeu 88
DB 182 AATGGCTAATGATGCAACCGCGTGAACACTTGTGGCCGAGAGGCTTCTGCAAGCTC 241
QY 89 LeuLeuGlyGlnLeuAsnAlaGlnHisAlaGluAlaAspValArgValLeuLeuGln 108
DB 242 CTCCTGGCCAGCTCAACACTGAGCACTGAGGCTGATGTGCGAGCGGCTCTGCTGACG 301
QY 109 AlaPheAspValAlaGlnArgSerPheLeuGlnSerTyrAspAspAlaLeuAlaGluLys 128
DB 302 GCCTTGAATGTGTGTGAGAGGAGCTTCCTGAGTCTATGATGATGATGCCCTTACTGAGAA 361
QY 129 AlaSerLeuGlnSerGlnLeuProGluGlyValProGlnHisGlnLeuProGlnTyr 148
DB 362 GCAGGCTCCAGTCCAGCTGCCGAGGGTGTGCCACACACAGCTGCACCTCAGTAC 421
QY 149 GlnLysIleLeuGlnArgLeuLysThrLeuGlnArgGlnIleSerGlyValAlaMetAla 168
DB 422 CAGAAATCTCTAGAGACTCAAGGCACTGAGAGGGAATTTCCGGAGAGACCATGGCT 481
QY 169 ValValAlaValLeuLeuAsnAsnLysLeuTyrValAlaAsnValGlyThrAsnArgAla 188
DB 482 GTCTGTGCAAGTCTCTCAACAGCAAGCTCTACGTTGCCAATGTGGTACGAACAGGGCC 541
QY 189 LeuLeuCysLysSerThrValAspGlyLeuGlnValThrGlnLeuAsnValAspHisThr 208
DB 542 CTTCTGTGCAAACTACAGTACAGTGGTTACAGGTTACACACTTAAACATGAGCAACAC 601
QY 209 Thr-GluAsnGluAspGlnLeuPheArgLeuSerGlnLeuGlyLeuAspAlaGlyLysIle 228
DB 602 ACCNGAAGACGAGAGAGAGCTCTTGGCTTCGCAACTGGGTTTGAAGACGAGAAAGAT 661
QY 228 eLysGlnValGlyIleIleCysGlyGlnLeuSerThrArgArgIleGlyAsp 245
DB 662 CAAGCAGATGGCGCTCATCTGTGACAGAGAGACAC-AGGCGGATTTGGGGAT 712
RESULT 14
AM247232 690 bp mRNA linear EST 07-JAN-2000
LOCUS 2820685.5prime NIH_MGC_7 Homo sapiens cDNA clone IMAGE:2820685 5',
DEFINITION mRNA sequence.
ACCESSION AM247232
VERSION AM247232.1 GI:6590225
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 690)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Other ESTs: 2820685.3prime
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: DCTD/DPF cDNA Library Preparation: Ling
Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E.
Consortium (LNL) DNA Sequencing by: Berkeley MGC sequencing
Project Clone distribution: MGC clone distribution information can
be found through the I.M.A.G.E. Consortium/LNL at:
www.bio.lnl.gov/bbrp/image/image.html Base Calling / Quality
Scores: PHRED from University of Washington Genome Center. Vector
Trimming: cross match from University of Washington Genome Center
PHRAP suite. Poly-T Identification: patmatch.pl from Berkeley
Drosophila Genome Project. University of Washington Genome Center:
http://www.genome.washington.edu
Place: LNCM4 row: N column: 14
High quality sequence stop: 513.

FEATURES
source Location/Qualifiers
1..690
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2820685"
/clone_1ib="NIH_MGC 7"
/rissue_type="small_cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dr priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(C) Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the Laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 160 a 182 c 217 g 131 t
ORIGIN
Alignment Scores:
Pred. No.: 6.9e-106 Length: 690
Score: 1135.00 Matches: 224
Percent Similarity: 99.13% Conservative: 3
Best Local Similarity: 97.82% Mismatches: 2
Query Match: 43.99% Indels: 0
DB: 10 Gaps: 0
US-09-830-144-4 (1-504) x AM247232 (1-690)
QY 133 SerGlnLeuProGluGlyValProGlnHisGlnLeuProProGlnTyrGlnLysIleLeu 152
DB 4 ACGAGTTGCCAGAGGAGTCCCTCAGACACACAGCTCCTCAGTATGAGAAAGATCCCTT 63
QY 153 GluArgLeuLysThrLeuGlnArgGlnIleSerGlyValAlaMetAlaValAlaVal 172
DB 64 GAGAGACTCAAGAGCTTACAGAGGGAATTTCCGGAGGGCCATGGCTGTGGCGGTC 123
QY 173 LeuLeuAsnAsnLysLeuTyrValAlaAsnValGlyThrAsnArgAlaLeuLeuCysLys 192
DB 124 CTTCTCAACAACAAGCTCTACGTGCGCAATGTGCGTACAAACCGTGCACTTTATGCAAA 183
QY 193 SerThrValAspGlyLeuGlnValThrGlnLeuAsnValAspHisThrThrGluAsnGlu 212
DB 184 TCGACAGTGTGATGGGTGTCAGGTGACACAGCTGACCTGACCAACCCACAGAAACGAG 243
QY 213 AspGlnLeuPheArgLeuSerGlnLeuGlyLeuAspAlaGlyLysIleLysGlnValGly 232
DB 244 GATGAGCTCTTCGCTCTTTGCGAGCTGGGCTTGGATGCTGGAAGATCAAGCAGGTGGG 303
QY 233 IleIleCysGlyGlnLeuSerThrArgArgIleGlyAspTyrLysValLysTyrGlyTyr 252
DB 304 ATCATCTGTGGGACGAGAGACACCCGCGGATTCGGGATTTACAAGTTAAATATGGCTAC 363
QY 253 ThrAspIleAspLeuLeuSerAlaAlaLysSerLysProIleIleAlaGluProGluIle 272
DB 364 ACGGACATTGACCTTCTTCACGCGCTGCCAATCCCAATCATCCACAGCCAGAAATC 423
QY 273 HisGlyValGlnProLeuAspGlyValThrGlyPheLeuValLeuMetSerGluGlyLeu 292
DB 424 CATGGGACAGCGCGCTGATGGGGTACGGGCTTCTGTGCTGATGTGGAGGGGTG 483
QY 293 TyrLysAlaLeuGlnAlaAlaHisGlyProGlyGlnAlaAsnGlnLysIleAlaMet 312
DB 484 TACAAGGCCCTAAGAGGAGCAGCCATGGGCTCTGTCAAGCCCAACGAGGATGCTGGGATG 543
QY 313 IleAspThrGluPheAlaLysGlnThrSerLeuAspAlaValAlaGlnAlaValAsp 332
DB 544 ATTGACACTGATGTTTGCACAGAGACCTCCCTGAGCGCACTGGCCCGAGCTGTGGAC 603
QY 333 ArgValLysArgIleHisSerAspThrPheAlaSerGlyGlyGlnArgAlaAspPheCys 352
DB 604 CGGGTAAAGCCCATTCACAGAGACCTTGCAGTGTGTGGAGGCTGCGCAAGTCTGC 663